

A priori error estimates of Crank-Nicolson finite element method for parabolic optimal control problems

Xindan Zhang^a, Jianping Zhao^{a,b,*}, Yanren Hou^{a,c}

^a College of Mathematics and System Sciences, Xinjiang University, Urumqi 830046, China

^b Institute of Mathematics and Physics, Xinjiang University, Urumqi 830046, China

^c School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China

ARTICLE INFO

Keywords:

Optimal control problem
Parabolic equation
Finite element method
Crank-Nicolson

ABSTRACT

In this paper, the fully discrete finite element approximation for parabolic optimal control problems with pointwise control constraints is studied. We use standard piecewise linear finite elements for the space discretization of the state, and Crank-Nicolson scheme for time discretization. For control discretization we consider piecewise linear finite elements approximation and variational discretization. We derive a priori error estimates for state, adjoint state and control. Finally, some numerical examples are provided to confirm our theoretical results.

1. Introduction

The optimal control of partial differential equations has a wide range of applications in engineering. More specifically, heat treatment in cancer treatment, optimal temperature control of concrete dams, optimization of oil production processes, temperature control, material processing and forming, chemical reaction, shape design, fluid control, and applications in aerospace, all involve solving optimal control problems described by partial differential equations. Therefore, it is particularly important to investigate efficient numerical methods for solving such problems. At present, many numerical methods can be used to solve these problems, such as finite element method, finite difference method, spectral method, etc. (see [2–6]).

To the best of our knowledge, the first contribution to the convergence analysis of elliptic optimal control problems was given in [7]. The authors studied the finite element discretizations of elliptic optimal control problems with pointwise inequality constraints on the control variables, and obtained the error estimates for the approximation of the optimal control and optimal state. In [8], a variational discretization method was proposed, which does not use explicit discretization of control variables but discretizes control variables implicitly by using the first-order optimality conditions and the discretization of the state and adjoint equations. Some superconvergence results of mixed finite element approximation for elliptic optimal control problems were obtained in [9,10,13]. The superconvergence of the finite element approximation for linear and semi-linear elliptic optimal control problems can be found in [18] and [12], respectively. A priori error estimates of the RT0 mixed finite element approximation for elliptic optimal control problems were established in [14]. The superconvergence of the linear and semi-linear elliptic optimal control problems with integral constraints has also been obtained [11,14].

Parabolic optimal control problems can often be described in environmental modelling, low-temperature superconducting laser energy blasting, petroleum reservoir simulations, and many other fields. The a priori error analysis for the finite element discretization of parabolic optimal control problems without control constraints was done in [2]. The a priori error estimate for space-time finite element discretizations of parabolic optimal control problems was derived in [3]. The Ritz-Galerkin approximation for parabolic optimal control problems was studied in [1]. The a posteriori error estimation of spectral method for parabolic optimal control problems was established in [6]. The Crank-Nicolson linear finite volume element method was considered for parabolic optimal control problems in [16]. In [23], the authors considered the Crank-Nicolson scheme for optimal control problems without control constraints. The superconvergence of the fully discrete finite element approximations for linear and semi-linear parabolic optimal control problems can be found in [19] and [20], respectively. The space-time finite element discretizations of time-optimal control

* Corresponding author at: College of Mathematics and System Sciences, Xinjiang University, Urumqi 830046, China.
E-mail address: zhaojianping@126.com (J. Zhao).

problems was studied in [15]. The superconvergence of fully discrete finite element for parabolic optimal control problems with integral constraints was obtained in [17]. Meidner and Vexler [22] considered a Petrov-Galerkin Crank-Nicolson scheme for parabolic optimal control problems where the controls are only time dependent. The temporal discretization is based on a Petrov-Galerkin variant of the Crank-Nicolson scheme, whereas the spatial discretization employs usual conforming finite elements. With a suitable postprocessing step on the dual partition, a second order convergence is obtained for the discrete control. Motivated by the work [22], the paper [26] considered the Crank-Nicolson time stepping and variational discretization for the similar problem as in [22]. The state equation is treated with a Petrov-Galerkin scheme using a piecewise constant ansatz for the state and piecewise linear continuous test functions. This results in variants of the Crank-Nicolson scheme for the state and the adjoint state. A second order convergence is obtained for the discrete control and the postprocessed state on the dual partition. Compared to the variational framework in [22] and [26], our result is based on the standard Crank-Nicolson difference scheme and provides a priori error estimate for such scheme. Thus, we need higher regularity requirements for the state and adjoint state.

In this paper, we study the fully discretization based on the Crank-Nicolson time discretization and linear finite element spatial discretization for parabolic optimal control problems. For the discretization of the state equation, we use standard piecewise linear finite elements in space and the Crank-Nicolson scheme in time. The discrete control variable is obtained by piecewise linear finite element or the projection of the discretized adjoint state on the set of admissible controls. For both cases, we derive an a priori error estimate for the state and control and then use numerical experiments to verify our theoretical results.

We consider the following optimal control problem:

$$\min_{u \in U_{ad}} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_{L^2(0,T;L^2(\Omega))}^2 \quad (1.1)$$

subject to

$$\begin{cases} y_t - \Delta y = f + Bu, & x \in \Omega, t \in (0, T], \\ y = 0, & x \in \Gamma, t \in (0, T], \\ y(\cdot, 0) = y_0, & x \in \Omega, \end{cases} \quad (1.2)$$

where $\Omega \subset \mathbb{R}^n (n = 2, 3)$ is a bounded convex polygon or polyhedron and Γ is the boundary of Ω . $\alpha > 0$, $T > 0$. We set the control space $U = L^2(0, T; L^2(\Omega))$. Let B be a continuous linear operator from $L^2(0, T; L^2(\Omega))$ to $L^2(0, T; L^2(\Omega))$, U_{ad} be a set defined by

$$U_{ad} = \{u : u \in U, u_a \leq u(x, t) \leq u_b, a.e.in \Omega \times (0, T)\},$$

where $u_a < u_b$ are two constants.

The rest of the paper is organized as follows. In section 2, we define the fully discrete finite-element approximation of parabolic optimal control problems and provide useful error estimates. In section 3, we present a priori error estimates for the finite element approximation and variational discretization of the optimal control problem. In section 4, we provide numerical examples to verify our theoretical results.

2. Optimal control problem

For $m \geq 0$ and $1 \leq p \leq \infty$, we use the standard Sobolev space $W^{m,p}(\Omega)$ with norm $\|\cdot\|_{m,p}$ and semi-norm $|\cdot|_{m,p}$ given by

$$\|v\|_{m,p}^p = \sum_{|\alpha| \leq m} \|D^\alpha v\|_{L^p}^p \quad \text{and} \quad |v|_{m,p}^p = \sum_{|\alpha|=m} \|D^\alpha v\|_{L^p}^p, \quad \forall v \in W^{m,p}(\Omega).$$

We denote by $W_0^{m,p}(\Omega)$ the closure of $C_0^\infty(\Omega)$ in $W^{m,p}(\Omega)$. For $p=2$, we denote $H^m(\Omega) = W^{m,2}(\Omega)$, $L^2(\Omega) = W^{0,2}(\Omega)$ and $H_0^m(\Omega) = W_0^{m,2}(\Omega)$.

We denote by $L^r(0, T; W^{m,p}(\Omega))$ the Banach space of all L^r integrable functions from $[0, T]$ to $W^{m,p}(\Omega)$ with the norm

$$\|v\|_{L^r(0,T;W^{m,p}(\Omega))} = \left(\int_0^T \|v\|_{m,p}^r dt \right)^{\frac{1}{r}} \quad \text{for } 1 \leq r < \infty,$$

and standard modification for $r = \infty$. We denote the $L^2(\Omega)$ -inner product by (\cdot, \cdot) .

Let

$$\begin{aligned} a(y, v) &= \int_{\Omega} \nabla y \cdot \nabla v \, dx, \quad \forall y, v \in H^1(\Omega), \\ (f_1, f_2) &= \int_{\Omega} f_1 f_2 \, dx, \quad \forall f_1, f_2 \in L^2(\Omega). \end{aligned}$$

Thus, a weak formulation for problem (1.1) is

$$\begin{cases} \min_{u \in U_{ad}} \frac{1}{2} \|y - y_d\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_{L^2(0,T;L^2(\Omega))}^2 \\ (y_t, w) + a(y, w) = (f + Bu, w), \quad \forall w \in H_0^1(\Omega), t \in (0, T], \\ y(\cdot, 0) = y_0, \quad x \in \Omega. \end{cases} \quad (2.1)$$

Assumption 2.1 ([22]). We assume that $f, y_d \in H^1(0, T; L^2(\Omega))$, $f(0), y_d(T) \in H_0^1(\Omega)$, and $\Delta y_0 \in H_0^1(\Omega)$.

Lemma 2.1 ([22]). For given $f, y_d \in L^2(0, T; L^2(\Omega))$, $y_0 \in H_0^1(\Omega)$ and $\alpha > 0$, there exists a unique solution (y, u) for problem (2.1).

The pair (y, u) is the solution to the problem (2.1) if and only if there is an adjoint state p such that the triplet (y, u, p) satisfies the following optimality conditions:

$$(y_t, w) + a(y, w) = (f + Bu, w), \quad \forall w \in H_0^1(\Omega), t \in (0, T], \\ y(\cdot, 0) = y_0, \quad x \in \Omega, \quad (2.2)$$

$$-(p_t, q) + a(p, q) = (y - y_d, q), \quad \forall q \in H_0^1(\Omega), t \in (0, T], \\ p(\cdot, T) = 0, \quad x \in \Omega, \quad (2.3)$$

$$(\alpha u + B^* p, v - u) \geq 0, \quad \forall v \in U_{ad}, t \in (0, T], \quad (2.4)$$

where B^* denotes the adjoint operator of B .

The optimality condition (2.4) can be equivalently formulated using the pointwise projection

$$P_{U_{ad}} : U \rightarrow U_{ad}, \quad P_{U_{ad}}(v)(x, t) = \max\{u_a, \min\{u_b, v(x, t)\}\}.$$

The resulting optimality condition reads

$$u = P_{U_{ad}} \left(-\frac{1}{\alpha} B^* p(u) \right). \quad (2.5)$$

It is well known that for $v \in L^2(0, T; W^{1,p}(\Omega))$ the projection $P_{U_{ad}}$ possesses the following property

$$\| \nabla(P_{U_{ad}}(v))(t) \|_{L^p(\Omega)} \leq \| \nabla v(t) \|_{L^p(\Omega)}. \quad (2.6)$$

Lemma 2.2 ([22]). *Let (y, u, p) be the solution of the optimal control problem (2.1)-(2.4). Then there holds the regularity:*

$$u \in L^2(0, T; W^{1,s}(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ y, p \in L^2(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega)),$$

for any $s < \infty$ when $n = 2$ and $s \leq 6$ when $n = 3$.

Furthermore, if Assumption 2.1 is valid, we have the improved regularity for the state and adjoint state, i.e.,

$$y, p \in H^1(0, T; H^2(\Omega)) \cap H^2(0, T; L^2(\Omega))$$

and the stability estimates

$$\begin{aligned} & \| y_{tt} \|_{L^2(0,T;L^2(\Omega))} + \| \Delta y \|_{H^1(0,T;L^2(\Omega))} \\ & \leq C(\| \nabla \Delta y_0 \|_{L^2(\Omega)} + \| \nabla f(0) \|_{L^2(\Omega)} + \| \nabla u(0) \|_{L^2(\Omega)} \\ & \quad + \| f \|_{H^1(0,T;L^2(\Omega))} + \| u \|_{H^1(0,T;L^2(\Omega))}), \\ & \| p_{tt} \|_{L^2(0,T;L^2(\Omega))} + \| \Delta p \|_{H^1(0,T;L^2(\Omega))} \\ & \leq C(\| y \|_{H^1(0,T;L^2(\Omega))} + \| y_d \|_{H^1(0,T;L^2(\Omega))} + \| \nabla y(T) \|_{L^2(\Omega)} + \| \nabla y_d(T) \|_{L^2(\Omega)}). \end{aligned}$$

Proof. The proof follows the idea of [22]. From [24, Proposition 2.1] we conclude that $y, p \in L^2(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$. The embedding $H^2(\Omega) \hookrightarrow W^{1,s}(\Omega)$ and (2.6) imply the result for u . Furthermore, if Assumption 2.1 is valid, we have that the right-hand side of the adjoint equation fulfills $y - y_d \in H^1(0, T; L^2(\Omega))$, and since $L^2(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega)) \hookrightarrow C(0, T; H_0^1(\Omega))$, we have that $y(T) - y_d(T) \in H_0^1(\Omega)$. We conclude that $p \in H^1(0, T; H^2(\Omega)) \cap H^2(0, T; L^2(\Omega))$ (see [24, Proposition 2.1]). By property (2.6), this implies $u(0) \in H^1(\Omega)$. Consequently, we obtain that $y \in H^1(0, T; H^2(\Omega)) \cap H^2(0, T; L^2(\Omega))$ (see [24, Proposition 2.1]). \square

Assumption 2.2. In the following analysis we require an additional regularity for y, p such that $y, p \in H^3(0, T; L^2(\Omega)) \cap H^2(0, T; H^2(\Omega))$. This regularity can be achieved if we remove the pointwise control constraints, or the controls are purely time dependent or spatial dependent, i.e., $Bu = u(t)g(x)$ or $Bu = u(x)g(t)$ for some given $g(x)$ or $g(t)$.

Let \mathcal{T}_h be a family of quasi-uniform partitions of Ω into disjoint regular simplexes K , such that $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} \bar{K}$. The finite element space V_h is defined as follows:

$$V_h = \{v \in C(\Omega) \cap H_0^1(\Omega) \mid v|_K \in \mathcal{P}_1(K), \forall K \in \mathcal{T}_h\},$$

where $\mathcal{P}_1(K)$ is the standard space of all polynomial functions with degrees not greater than 1 on K .

To obtain an optimal error estimate, we introduce the following projection operators. We define the Ritz projection $R_h : H_0^1(\Omega) \rightarrow V_h$ as:

$$(\nabla R_h u, \nabla v_h) = (\nabla u, \nabla v_h), \quad \forall v_h \in V_h,$$

and the L^2 projection $P_h : L^2(\Omega) \rightarrow V_h$ as:

$$(P_h u, v_h) = (u, v_h), \quad \forall v_h \in V_h.$$

We also define the d Lagrange interpolation $I_d : C(\Omega) \rightarrow V_h$ as:

$$(I_d v)(p_i) = v(p_i) \text{ for each node } p_i \text{ of } \mathcal{T}_h.$$

It is well known that the Ritz projection satisfies [21]

$$\| w - R_h w \|_{L^2(\Omega)} \leq C h^2 \| w \|_{H^2(\Omega)}.$$

For a given integer N , let $\Delta t = T/N$ be the uniform time step size and the nodes are denoted by $t_n = n\Delta t$, $0 \leq n \leq N$. Let $v^i = v(x, t_i)$ and $d_t v^i = (v^i - v^{i-1})/\Delta t$. We define a discrete time-dependent norm as

$$|||v||| = \left(\sum_{n=1}^N \Delta t \left\| \frac{v^n + v^{n-1}}{2} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}.$$

We present the Crank-Nicolson linear finite-element method for the former system: Find $(Y_h^n, P_h^n, U_h^n) \in V_h \times V_h \times U_{ad,h}$ such that for all $w_h \in V_h, q_h \in V_h, v \in U_{ad,h}$,

$$\left(\frac{Y_h^n - Y_h^{n-1}}{\Delta t}, w_h \right) + a\left(\frac{Y_h^n + Y_h^{n-1}}{2}, w_h \right) = \left(\frac{f^n + BU_h^n + f^{n-1} + BU_h^{n-1}}{2}, w_h \right), \quad (2.7)$$

$$Y_h^0(x) = R_h y_0, \quad x \in \Omega, \quad \forall w_h \in V_h, n = 1, 2, \dots, N.$$

$$\left(\frac{P_h^n - P_h^{n-1}}{\Delta t}, q_h \right) + a\left(\frac{P_h^n + P_h^{n-1}}{2}, q_h \right) = \left(\frac{Y_h^n - y_d^n + Y_h^{n-1} - y_d^{n-1}}{2}, q_h \right), \quad (2.8)$$

$$P_h^N(x) = 0, \quad x \in \Omega, \quad \forall q_h \in V_h, n = 1, 2, \dots, N.$$

$$\left(\alpha \frac{U_h^n + U_h^{n-1}}{2} + B^* \frac{P_h^n + P_h^{n-1}}{2}, v - \frac{U_h^n + U_h^{n-1}}{2} \right) \geq 0, \quad \forall v \in U_{ad,h}, n = 1, 2, \dots, N. \quad (2.9)$$

In the following we denote Y_h , P_h and U_h as fully discrete finite element approximations of y , p and u , respectively. Here, $U_{ad,h}$ is an appropriate approximation of U_{ad} . We set $U_{ad,h} \equiv V_h \cap U_{ad}$ for piecewise linear discretization of the control problem and $U_{ad,h} \equiv U_{ad}$ for its variational discretization.

3. Error estimates for the optimal control problems

Throughout this section we will assume the solutions y, p possess the regularity $y, p \in H^3(0, T; L^2(\Omega) \cap H^2(0, T; H^2(\Omega)))$. For such a regularity we need smooth right-hand sides and additional compatibility relations.

First, we define some intermediate variables. For any $v \in U_{ad}$, let $(y(v), p(v))$ be the solution to the following equations:

$$(y_t(v), w) + a(y(v), w) = (f + Bv, w), \quad \forall w \in H_0^1(\Omega), t \in (0, T], \quad (3.1)$$

$$y(v)(\cdot, 0) = y_0, \quad x \in \Omega.$$

$$-(p_t(v), q) + a(p(v), q) = (y(v) - y_d, q), \quad \forall q \in H_0^1(\Omega), t \in (0, T], \quad (3.2)$$

$$p(v)(\cdot, T) = 0, \quad x \in \Omega.$$

For any $v \in U_{ad}$, let $(y_h(v), p_h(v))$ be the solution to the following equations:

$$(y_{ht}(v), w_h) + a(y_h(v), w_h) = (f + Bv, w_h), \quad \forall w_h \in V_h, t \in (0, T], \quad (3.3)$$

$$y_h(v)(\cdot, 0) = R_h y_0, \quad x \in \Omega.$$

$$-(p_{ht}(v), q_h) + a(p_h(v), q_h) = (y_h(v) - y_d, q_h), \quad \forall q_h \in V_h, t \in (0, T], \quad (3.4)$$

$$p_h(v)(\cdot, T) = 0, \quad x \in \Omega.$$

For any $v \in U_{ad}$, the pair $(y_h^n(v), p_h^n(v))$ for $n = 1, 2, \dots, N$ satisfies the following system:

$$\left(\frac{y_h^n(v) - y_h^{n-1}(v)}{\Delta t}, w_h \right) + a\left(\frac{y_h^n(v) + y_h^{n-1}(v)}{2}, w_h \right) = \left(\frac{f^n + Bv^n + f^{n-1} + Bv^{n-1}}{2}, w_h \right),$$

$$y_h^0(v)(x) = R_h y_0, \quad x \in \Omega, \quad \forall w_h \in V_h, n = 1, 2, \dots, N. \quad (3.5)$$

$$\left(\frac{p_h^n(v) - p_h^{n-1}(v)}{\Delta t}, q_h \right) + a\left(\frac{p_h^n(v) + p_h^{n-1}(v)}{2}, q_h \right) = \left(\frac{y_h^n(v) - y_d^n + y_h^{n-1}(v) - y_d^{n-1}}{2}, q_h \right),$$

$$p_h^N(v) = 0, \quad x \in \Omega, \quad \forall q_h \in V_h, n = 1, 2, \dots, N. \quad (3.6)$$

For the following error estimate, we introduce the Gronwall lemma [24] and discrete Gronwall inequality [25].

Lemma 3.1 (Gronwall lemma). Let f be a non-negative function and let g and φ be continuous functions on $[t_0, T]$. If φ satisfies

$$\varphi(t) \leq g(t) + \int_{t_0}^t f(\tau) \varphi(\tau) d\tau, \quad \forall t \in [t_0, T],$$

then

$$\varphi(t) \leq g(t) + \int_{t_0}^t f(s) g(s) \exp\left(\int_s^t f(\tau) d\tau\right) ds, \quad \forall t \in [t_0, T].$$

If g is non-decreasing, then

$$\varphi(t) \leq g(t) \exp\left(\int_{t_0}^t f(\tau) d\tau\right), \quad \forall t \in [t_0, T].$$

Lemma 3.2 (Discrete Gronwall lemma). Let $k, B, a_n, b_n, c_n, \alpha_n$ be non-negative numbers for integers $n \geq 1$ and let the inequality

$$a_{N+1} + k \sum_{n=1}^{N+1} b_n \leq B + k \sum_{n=1}^{N+1} c_n + k \sum_{n=1}^{N+1} \alpha_n a_n \quad \text{for } N \geq 0$$

hold. If $k\alpha_n < 1$ for all $n = 1, 2, \dots, N+1$, then

$$a_{N+1} + k \sum_{n=1}^{N+1} b_n \leq (B + k \sum_{n=1}^{N+1} c_n) \exp(k \sum_{n=1}^{N+1} \frac{\alpha_n}{1 - k\alpha_n}) \quad \text{for } N \geq 0.$$

If the inequality

$$a_{N+1} + k \sum_{n=1}^{N+1} b_n \leq B + k \sum_{n=1}^{N+1} c_n + k \sum_{n=1}^{N+1} \alpha_n a_n \quad \text{for } N \geq 0$$

is given, then it holds

$$a_{N+1} + k \sum_{n=1}^{N+1} b_n \leq (B + k \sum_{n=1}^{N+1} c_n) \exp(k \sum_{n=1}^{N+1} \alpha_n) \quad \text{for } N \geq 0.$$

Theorem 3.1. Let $y(u) \in L^2(0, T; H^2(\Omega))$ be the solution to problem (3.1), satisfying $y_t(u) \in L^2(0, T; H^2(\Omega))$ and $y_h(u) \in V_h$ be the solution to problem (3.3). Subsequently, we have

$$\|y(u) - y_h(u)\|_{L^2(\Omega)}^2 \leq C(T)h^4 \left(\|y(u)\|_{H^2(\Omega)}^2 + \int_0^T \|y_t(u)\|_{H^2(\Omega)}^2 dt \right).$$

Proof. From (3.1) and (3.3), we obtain

$$(y_t(u) - y_{ht}(u), w_h) + a(y(u) - y_h(u), w_h) = 0.$$

Using the Ritz projection, we have

$$(y_t(u) - R_h y_t(u), w_h) + (R_h y_t(u) - y_{ht}(u), w_h) + a(R_h y(u) - y_h(u), w_h) = 0. \quad (3.7)$$

Setting $e = R_h y(u) - y_h(u)$, taking $w_h = R_h y(u) - y_h(u)$ in (3.7) and using the definition of R_h and Young's inequality, we obtain

$$\frac{d}{dt} \|e\|_{L^2(\Omega)}^2 + \|\nabla e\|_{L^2(\Omega)}^2 \leq c_0 \|e\|_{L^2(\Omega)}^2 + c_0 \|y_t(u) - R_h y_t(u)\|_{L^2(\Omega)}^2.$$

From the properties of R_h , we have that

$$\frac{d}{dt} \|e\|_{L^2(\Omega)}^2 + \|\nabla e\|_{L^2(\Omega)}^2 \leq c_0 \|e\|_{L^2(\Omega)}^2 + c_1 h^4 \|y_t(u)\|_{H^2(\Omega)}^2. \quad (3.8)$$

Integrating (3.8) from 0 to t and using Lemma 3.1, we obtain

$$\|e\|_{L^2(\Omega)}^2 + \int_0^t \|\nabla e\|_{L^2(\Omega)}^2 dt \leq c_1 h^4 \exp(c_0 T) \int_0^T \|y_t(u)\|_{H^2(\Omega)}^2 dt. \quad (3.9)$$

Using (3.9) and the definition of R_h , we deduce that

$$\begin{aligned} \|y(u) - y_h(u)\|_{L^2(\Omega)}^2 &\leq \|y(u) - R_h y(u)\|_{L^2(\Omega)}^2 + \|R_h y(u) - y_h(u)\|_{L^2(\Omega)}^2 \\ &\leq Ch^4 \|y(u)\|_{H^2(\Omega)}^2 + C(T)h^4 \int_0^T \|y_t(u)\|_{H^2(\Omega)}^2 dt. \quad \square \end{aligned}$$

Theorem 3.2. Assume that h and Δt are both sufficiently small, $y(u)$ is the solution of (3.1), and $y_h(u)$ is the solution of (3.3), satisfying $y_{htt}(u) \in L^2(0, T; L^2(\Omega))$. The solution sequence $\{y_h^n(u)\}$ of (3.5) satisfies the following error estimate

$$\|y_h(u)(t_n) - y_h^n(u)\|_{L^2(\Omega)}^2 \leq C(T)(\Delta t)^4 \int_0^T \|y_{htt}(u)\|_{L^2(\Omega)}^2 dt. \quad (3.10)$$

Proof. We prove (3.10) using the mathematical induction method. First, (3.10) holds true for $n = 0$. We assume that (3.10) is true for $n = k - 1$. We must prove (3.10) for $n = k$.

It follows (3.3) with $t = t_n$ and $t = t_{n-1}$ that

$$\begin{aligned} & \left(\frac{y_{ht}(u)(t_n) + y_{ht}(u)(t_{n-1})}{2}, w_h \right) + a\left(\frac{y_h(u)(t_n) + y_h(u)(t_{n-1})}{2}, w_h \right) \\ &= \left(\frac{f^n + Bu^n + f^{n-1} + Bu^{n-1}}{2}, w_h \right). \end{aligned} \quad (3.11)$$

From (3.11) and (3.5), we obtain

$$\begin{aligned} & \left(\frac{y_{ht}(u)(t_n) + y_{ht}(u)(t_{n-1})}{2} - d_t y_h^n(u), w_h \right) \\ &+ a\left(\frac{y_h(u)(t_n) + y_h(u)(t_{n-1}) - y_h^{n-1}(u) - y_h^n(u)}{2}, w_h \right) = 0. \end{aligned} \quad (3.12)$$

Setting $e_h^n = y_h(u)(t_n) - y_h^n(u)$ and $w_h = \bar{e}_h^n = \frac{e_h^n + e_h^{n-1}}{2}$ in equation (3.12), we obtain

$$\begin{aligned} & \frac{1}{2\Delta t} \left(\| e_h^n \|_{L^2(\Omega)}^2 - \| e_h^{n-1} \|_{L^2(\Omega)}^2 \right) + a(\bar{e}_h^n, \bar{e}_h^n) \\ & \leq C \| d_t y_h(u)(t_n) - \frac{y_{ht}(u)(t_n) + y_{ht}(u)(t_{n-1})}{2} \|_{L^2(\Omega)} \left(\| e_h^n \|_{L^2(\Omega)} + \| e_h^{n-1} \|_{L^2(\Omega)} \right) \\ & \leq C_1(\Delta t)^{-1} \left\| \int_{t_{n-1}}^{t_n} (t - t_{n-1})(t_n - t) y_{htt}(u) dt \right\|_{L^2(\Omega)} \left(\| e_h^n \|_{L^2(\Omega)} + \| e_h^{n-1} \|_{L^2(\Omega)} \right) \\ & \leq C_2(\Delta t)^{\frac{3}{2}} \left(\int_{t_{n-1}}^{t_n} \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}} \left(\| e_h^n \|_{L^2(\Omega)} + \| e_h^{n-1} \|_{L^2(\Omega)} \right) \\ & \leq C(\delta)(\Delta t)^3 \int_{t_{n-1}}^{t_n} \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt + \delta \left(\| e_h^n \|_{L^2(\Omega)}^2 + \| e_h^{n-1} \|_{L^2(\Omega)}^2 \right). \end{aligned}$$

Hence, we have

$$\begin{aligned} & \| e_h^n \|_{L^2(\Omega)}^2 - \| e_h^{n-1} \|_{L^2(\Omega)}^2 + c_1 \Delta t \| \nabla \bar{e}_h^n \|_{L^2(\Omega)}^2 \\ & \leq C(\delta)(\Delta t)^4 \int_{t_{n-1}}^{t_n} \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt + \delta \Delta t \left(\| e_h^n \|_{L^2(\Omega)}^2 + \| e_h^{n-1} \|_{L^2(\Omega)}^2 \right). \end{aligned} \quad (3.13)$$

Summing (3.13) from $n = 1$ to $n = k$, we obtain

$$\begin{aligned} & \| e_h^k \|_{L^2(\Omega)}^2 + c_1 \Delta t \sum_{n=1}^k \| \nabla \bar{e}_h^n \|_{L^2(\Omega)}^2 \\ & \leq C(\delta)(\Delta t)^4 \int_0^T \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt + 2\delta \Delta t \sum_{n=1}^k \| e_h^n \|_{L^2(\Omega)}^2. \end{aligned} \quad (3.14)$$

Based on Lemma 3.2, we get the estimates

$$\| e_h^k \|_{L^2(\Omega)}^2 + c_1 \Delta t \sum_{n=1}^k \| \nabla \bar{e}_h^n \|_{L^2(\Omega)}^2 \leq \exp(4\delta T) C(\delta)(\Delta t)^4 \int_0^T \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt. \quad \square \quad (3.15)$$

Theorem 3.3. Assume that the conditions in Theorem 3.1 and Theorem 3.2 hold, then y^n and $y_h^n(u)$ satisfy the following error estimate:

$$\begin{aligned} & \| y^n(u) - y_h^n(u) \|_{L^2(\Omega)}^2 \leq C(T)(\Delta t)^4 \int_0^T \| y_{htt}(u) \|_{L^2(\Omega)}^2 dt \\ & + C_1(T)h^4 \left(\| y(u) \|_{H^2(\Omega)}^2 + \int_0^T \| y_t(u) \|_{H^2(\Omega)}^2 dt \right). \end{aligned} \quad (3.16)$$

Proof. Combining Theorem 3.1 with Theorem 3.2 yields the error estimate. \square

Theorem 3.4. Under the assumption of Theorem 3.3, let $p(u)$, $p_h(u)$ and $p_h^n(u)$ be the solutions of (3.2), (3.4), and (3.6), respectively, and

$$p(u), p_t(u) \in L^2(0, T; H^2(\Omega)), p_{htt}(u) \in L^2(0, T; L^2(\Omega)).$$

We have

$$\begin{aligned} \| p^n(u) - p_h^n(u) \|_{L^2(\Omega)}^2 &\leq C(T)(\Delta t)^4 \left(\int_0^T \| p_{hhtt}(u) \|_{L^2(\Omega)}^2 dt + \Delta t \int_0^T \| y_{hhtt}(u) \|_{L^2(\Omega)}^2 dt \right) \\ &\quad + C_1(T)h^4 \left(\| y(u) \|_{H^2(\Omega)}^2 + \int_0^T \| y_t(u) \|_{H^2(\Omega)}^2 dt + \| p(u) \|_{H^2(\Omega)}^2 + \int_0^T \| p_t(u) \|_{H^2(\Omega)}^2 dt \right). \end{aligned}$$

Proof. The proof is similar to that of Theorem 3.1, so we omit it. \square

Theorem 3.5. Let (y_h^n, p_h^n) and (Y_h^n, P_h^n) be the solutions to problem (3.5)-(3.6) and (2.5)-(2.6), respectively. Then

$$\begin{aligned} \| y_h^n(u) - Y_h^n \|_{H^1(\Omega)} &\leq C ||| u - U_h |||, \\ \| p_h^n(u) - P_h^n \|_{H^1(\Omega)} &\leq C ||| u - U_h |||. \end{aligned}$$

Proof. From (3.5) and (2.7), we obtain

$$(d_t(Y_h^i - y_h^i(u)), w_h) + a\left(\frac{Y_h^i - y_h^i(u) + Y_h^{i-1} - y_h^{i-1}(u)}{2}, w_h\right) = \left(B\left(\frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2}\right), w_h\right). \quad (3.17)$$

Setting $\eta^i = Y_h^i - y_h^i(u)$ and $w_h = \frac{\eta^i - \eta^{i-1}}{\Delta t}$ in equation (3.17), we have

$$(d_t\eta^i, d_t\eta^i) + a\left(\frac{\eta^i + \eta^{i-1}}{2}, \frac{\eta^i - \eta^{i-1}}{\Delta t}\right) = \left(B\left(\frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2}\right), \frac{\eta^i - \eta^{i-1}}{\Delta t}\right). \quad (3.18)$$

We obtain

$$(d_t\eta^i, d_t\eta^i) + \frac{1}{2\Delta t} [a(\eta^i, \eta^i) - a(\eta^{i-1}, \eta^{i-1})] = \left(B\left(\frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2}\right), d_t\eta^i\right). \quad (3.19)$$

The property of continuity for B implies that

$$\left(B\left(\frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2}\right), d_t\eta^i\right) \leq C(\delta) \left\| \frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2} \right\|_{L^2(\Omega)}^2 + \delta \| d_t\eta^i \|_{L^2(\Omega)}^2. \quad (3.20)$$

Summing i from 1 to n and noting $\eta^0 = 0$, we obtain

$$\| \eta^n \|_{H^1(\Omega)}^2 \leq C \sum_{i=1}^N \Delta t \left\| \frac{U_h^i - u^i + U_h^{i-1} - u^{i-1}}{2} \right\|_{L^2(\Omega)}^2 = C ||| u - U_h |||^2. \quad (3.21)$$

Let $e^i = P_h^i - p_h^i(u)$. Similarly, we obtain

$$a(e^{i-1}, e^{i-1}) \leq a(e^i, e^i) + C \triangle t \left\| \frac{Y_h^i - y_h^i(u) + Y_h^{i-1} - y_h^{i-1}(u)}{2} \right\|_{L^2(\Omega)}^2.$$

Summing i from $n+1$ to N and noting $e^N = 0$, we have

$$a(e^n, e^n) \leq C \sum_{i=1}^N \Delta t \left\| \frac{Y_h^i - y_h^i(u) + Y_h^{i-1} - y_h^{i-1}(u)}{2} \right\|_{L^2(\Omega)}^2,$$

we obtain

$$\| P_h^n - p_h^n(u) \|_{H^1(\Omega)} \leq C ||| u - U_h |||. \quad \square$$

3.1. Variational discretization

As first, we consider the variational control discretization introduced in [8], i.e., $U_{ad,h} \equiv U_{ad}$.

Theorem 3.6. Let (y, p, u) and (Y_h, P_h, U_h) be the solutions of problems (2.2)-(2.4) and (2.5)-(2.7), respectively. If the assumptions of Theorem 3.4 hold, we have

$$||| u - U_h ||| \leq C(h^2 + \Delta t^2).$$

Proof. From (2.4) and (2.9), we obtain

$$\begin{aligned} \alpha ||| u - U_h ||| ^2 &= \sum_{i=1}^N \alpha \Delta t \left(\frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2}, \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &= \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i + u^{i-1}}{2} + B^* \left(\frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &\quad - \sum_{i=1}^N \Delta t \left(\alpha \frac{U_h^i + U_h^{i-1}}{2} + B^* \left(\frac{P_h^i + P_h^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^N \Delta t \left(B^* \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\
& \leq \sum_{i=1}^N \Delta t \left(B^* \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\
& = T_1.
\end{aligned}$$

Let $p_h^i(u)$ be the solution of (3.6) satisfying the assumption of Theorem 3.4. Using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
T_1 & = \sum_{i=1}^N \Delta t \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2}, B \left(\frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \right) \right) \\
& = \sum_{i=1}^N \Delta t \left(\frac{p_h^i(u) + p_h^{i-1}(u)}{2} - \frac{p^i + p^{i-1}}{2}, B \left(\frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \right) \right) \\
& \quad + \sum_{i=1}^N \Delta t \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p_h^i(u) + p_h^{i-1}(u)}{2}, B \left(\frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \right) \right) \\
& = J_1 + J_2.
\end{aligned}$$

For the first term J_1 , using Theorem 3.4 we deduce that

$$\begin{aligned}
J_1 & = \sum_{i=1}^N \Delta t \left(\frac{p_h^i(u) + p_h^{i-1}(u)}{2} - \frac{p^i + p^{i-1}}{2}, B \left(\frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \right) \right) \\
& \leq C \sum_{i=1}^N \Delta t \| p_h^i(u) - p^i \|_{L^2(\Omega)} \| \frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \|_{L^2(\Omega)} \\
& \quad + C \sum_{i=1}^N \Delta t \| p_h^{i-1}(u) - p^{i-1} \|_{L^2(\Omega)} \| \frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \|_{L^2(\Omega)} \\
& \leq C \left(\sum_{i=1}^N \Delta t \| p_h^i(u) - p^i \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^N \Delta t \| \frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \\
& \quad + C \left(\sum_{i=1}^N \Delta t \| p_h^{i-1}(u) - p^{i-1} \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^N \Delta t \| \frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \\
& \leq C(h^2 + \Delta t^2) ||| u - U_h |||.
\end{aligned}$$

Let

$$\gamma^i = \frac{P_h^i + P_h^{i-1}}{2} - \frac{p_h^i(u) + p_h^{i-1}(u)}{2} \text{ and } \delta^i = \frac{y_h^i(u) + y_h^{i-1}(u)}{2} - \frac{Y_h^i + Y_h^{i-1}}{2},$$

we have

$$\begin{aligned}
J_2 & = \sum_{i=1}^N \Delta t \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p_h^i(u) + p_h^{i-1}(u)}{2}, B \left(\frac{u^i - U_h^i + u^{i-1} - U_h^{i-1}}{2} \right) \right) \\
& = \sum_{i=1}^N \Delta t (d_t(y_h^i(u) - Y_h^i), \gamma^i) + \sum_{i=1}^N \Delta t a(\gamma^i, \delta^i) \\
& = \sum_{i=1}^N \Delta t (d_t(y_h^i(u) - Y_h^i), \gamma^i) + \sum_{i=1}^N \Delta t (d_t(P_h^i - p_h^i(u)), \delta^i) \\
& \quad + \sum_{i=1}^N \Delta t \left(\frac{Y_h^i + Y_h^{i-1}}{2} - \frac{y_h^i(u) + y_h^{i-1}(u)}{2}, \delta^i \right) \\
& = R_1 + R_2 + R_3.
\end{aligned}$$

For terms R_1 and R_2 , noting $P_h^N - p_h^N(u) = 0$ and $Y_h^0 - y_h^0(u) = 0$, we obtain

$$\begin{aligned}
R_1 + R_2 & = \sum_{i=1}^N \Delta t (d_t(y_h^i(u) - Y_h^i), \gamma^i) + \sum_{i=1}^N \Delta t (d_t(P_h^i - p_h^i(u)), \delta^i) \\
& = \sum_{i=1}^N (y_h^i(u) - Y_h^i, P_h^i - p_h^i(u)) - \sum_{i=1}^N (y_h^{i-1}(u) - Y_h^{i-1}, P_h^{i-1} - p_h^{i-1}(u)) \\
& = 0.
\end{aligned}$$

By noting that $R_3 \leq 0$, we obtain $J_2 \leq 0$.

From the estimates of the terms T_1 , T_2 , J_1 and J_2 , we have

$$||| u - U_h ||| \leq C(h^2 + \Delta t^2). \quad \square$$

Combining Theorem 3.5 with Theorem 3.6 yields the following error estimation:

Theorem 3.7. Let (y, p, u) and (Y_h, P_h, U_h) be the solutions of problems (2.2)-(2.4), and (2.7)-(2.9), respectively. Then, we obtain the a priori error estimate:

$$\|y^n - Y_h^n\|_{L^2(\Omega)} + \|p^n - P_h^n\|_{L^2(\Omega)} \leq C(h^2 + k^2), \quad 0 \leq n \leq N. \quad (3.22)$$

Proof. Let $p_h^n(u)$ and $y_h^n(u)$ be the solutions of (3.6) and (3.5), respectively, satisfying the assumption of Theorem 3.4. Using the triangle inequality, we have

$$\begin{aligned} \|y^n - Y_h^n\|_{L^2(\Omega)} &\leq \|y^n - y_h^n(u)\|_{L^2(\Omega)} + \|y_h^n(u) - Y_h^n\|_{L^2(\Omega)}, \\ \|p^n - P_h^n\|_{L^2(\Omega)} &\leq \|p^n - p_h^n(u)\|_{L^2(\Omega)} + \|p_h^n(u) - P_h^n\|_{L^2(\Omega)}. \end{aligned}$$

Using Theorem 3.5, we obtain

$$\begin{aligned} \|y^n - Y_h^n\|_{L^2(\Omega)} &\leq \|y^n - y_h^n(u)\|_{L^2(\Omega)} + C|||u - U_h|||, \\ \|p^n - P_h^n\|_{L^2(\Omega)} &\leq \|p^n - p_h^n(u)\|_{L^2(\Omega)} + C|||u - U_h|||. \end{aligned}$$

Therefore, we obtain (3.22) from Theorems 3.3, 3.4, and 3.6. \square

3.2. Piecewise linear control discretization

We now consider the fully discrete case where the discretized state and adjoint are approximated by piecewise linear elements and the discrete control is searched for in $U_{ad,h} = V_h \cap U_{ad}$.

To derive an a priori error estimate we make the following assumption regarding the continuous control variable: For each time interval I_n we grouped mesh T_h into three sets $T_{h,n}^1 \cup T_{h,n}^2 \cup T_{h,n}^3$ with $T_{h,n}^i \cap T_{h,n}^j = \emptyset$ for $i \neq j$. The sets are defined as follows:

$$\begin{aligned} T_{h,n}^1 &:= \{K \in T_h | u(t_n, x) = u_a \text{ or } u(t_n, x) = u_b, \forall x \in K\}, \\ T_{h,n}^2 &:= \{K \in T_h | u_a < u(t_n, x) < u_b, \forall x \in K\}, \\ T_{h,n}^3 &:= T_h \setminus (T_{h,n}^1 \cup T_{h,n}^2). \end{aligned}$$

Assumption 3.1. There exists a positive constant C that is independent of $\Delta t, h$, and n such that

$$\sum_{K \in T_{h,n}^3} |K| \leq Ch.$$

A similar assumption is used in [3].

Theorem 3.8. Let (y, p, u) and (Y_h, P_h, U_h) be the solutions of problems (2.2)-(2.4) and (2.7)-(2.9), respectively. If the assumptions of Theorem 3.4 hold, we have

$$|||u - U_h||| \leq C(h^{\frac{3}{2}-\frac{1}{s}} + \Delta t^2).$$

Proof. From (2.4) and (2.9), we obtain

$$\begin{aligned} \alpha |||u - U_h|||^2 &= \sum_{i=1}^N \alpha \Delta t \left(\frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2}, \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &= \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i + u^{i-1}}{2} + B^* \left(\frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &\quad - \sum_{i=1}^N \Delta t \left(\alpha \frac{U_h^i + U_h^{i-1}}{2} + B^* \left(\frac{P_h^i + P_h^{i-1}}{2} \right), \frac{I_d u^i + I_d u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &\quad - \sum_{i=1}^N \Delta t \left(\alpha \frac{U_h^i + U_h^{i-1}}{2} + B^* \left(\frac{P_h^i + P_h^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &\quad + \sum_{i=1}^N \Delta t \left(B^* \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &\leq \sum_{i=1}^N \Delta t \left(B^* \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{U_h^i + U_h^{i-1}}{2} \right) \\ &\quad - \sum_{i=1}^N \Delta t \left(\alpha \frac{U_h^i + U_h^{i-1}}{2} + B^* \left(\frac{P_h^i + P_h^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &= T_1 + T_2. \end{aligned}$$

For the second term T_2 , we derive that

$$\begin{aligned} T_2 &= -\sum_{i=1}^N \Delta t \left(\alpha \frac{U_h^i + U_h^{i-1}}{2} + B^* \left(\frac{P_h^i + P_h^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &= \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i - U_h^i}{2} + \frac{u^{i-1} - U_h^{i-1}}{2}, \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &\quad + \sum_{i=1}^N \Delta t \left(B^* \left(\frac{p^i + p^{i-1} - P_h^i - P_h^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &\quad - \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i + u^{i-1}}{2} + B^* \left(\frac{p^i + p^{i-1}}{2} \right), \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &= J_1 + J_2 + J_3. \end{aligned}$$

Application of the Cauchy-Schwarz inequality gives

$$\begin{aligned} J_1 &= \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i - U_h^i}{2} + \frac{u^{i-1} - U_h^{i-1}}{2}, \frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \\ &\leq C \sum_{i=1}^N \Delta t \left\| \frac{u^i - U_h^i}{2} + \frac{u^{i-1} - U_h^{i-1}}{2} \right\|_{L^2(\Omega)} \left(\|u^i - I_d u^i\|_{L^2(\Omega)} + \|u^{i-1} - I_d u^{i-1}\|_{L^2(\Omega)} \right) \\ &\leq C \|u - U_h\| \left[\left(\sum_{i=1}^N \Delta t \|u^i - I_d u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^N \Delta t \|u^{i-1} - I_d u^{i-1}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \right]. \end{aligned}$$

Let $p_h^i(u)$ be the solution of (3.6) satisfying the assumption of Theorem 3.4. Using the Cauchy-Schwarz inequality, Theorem 3.2 and Theorem 3.5, we have

$$\begin{aligned} J_2 &= \sum_{i=1}^N \Delta t \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p^i + p^{i-1}}{2}, B \left(\frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \right) \\ &= \sum_{i=1}^N \Delta t \left(\frac{P_h^i + P_h^{i-1}}{2} - \frac{p_h^i(u) + p_h^{i-1}(u)}{2}, B \left(\frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \right) \\ &\quad + \sum_{i=1}^N \Delta t \left(\frac{p_h^i(u) + p_h^{i-1}(u)}{2} - \frac{p^i + p^{i-1}}{2}, B \left(\frac{u^i + u^{i-1}}{2} - \frac{I_d u^i + I_d u^{i-1}}{2} \right) \right) \\ &\leq C(h^2 + k^2) \left[\left(\sum_{i=1}^N \Delta t \|u^i - I_d u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^N \Delta t \|u^{i-1} - I_d u^{i-1}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \right] \\ &\quad + C \|u - U_h\| \left[\left(\sum_{i=1}^N \Delta t \|u^i - I_d u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^N \Delta t \|u^{i-1} - I_d u^{i-1}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \right]. \end{aligned}$$

For the third term J_3 , we derive that

$$\begin{aligned} J_3 &= \sum_{i=1}^N \Delta t \left(\alpha \frac{u^i + u^{i-1}}{2} + B^* \left(\frac{p^i + p^{i-1}}{2} \right), \frac{I_d u^i + I_d u^{i-1}}{2} - \frac{u^i + u^{i-1}}{2} \right) \\ &= \sum_{i=1}^N \Delta t \left(\alpha u^i + B^* p^i, \frac{I_d u^i - u^i}{2} \right) \\ &\quad + \sum_{i=1}^N \Delta t \left(\alpha u^{i-1} + B^* p^{i-1}, \frac{I_d u^{i-1} - u^{i-1}}{2} \right) \\ &\quad + \sum_{i=0}^N \Delta t \left(\alpha \frac{u^{i+1} + u^{i-1}}{2} + B^* \left(\frac{p^{i+1} + p^{i-1}}{2} \right) - \alpha u^i - B^* p^i, \frac{I_d u^i - u^i}{2} \right) \\ &\leq \sum_{i=1}^N \Delta t \left| \left(\alpha u^i + B^* p^i, \frac{I_d u^i - u^i}{2} \right) \right| \\ &\quad + \sum_{i=1}^N \Delta t \left| \left(\alpha u^{i-1} + B^* p^{i-1}, \frac{I_d u^{i-1} - u^{i-1}}{2} \right) \right| \\ &\quad + \sum_{i=0}^N \Delta t \left| \left(\alpha \frac{u^{i+1} + u^{i-1}}{2} + B^* \left(\frac{p^{i+1} + p^{i-1}}{2} \right) - \alpha u^i - B^* p^i, \frac{I_d u^i - u^i}{2} \right) \right| \\ &\quad + \sum_{i=0}^N \Delta t \left| \left(B^* \left(\frac{p^{i+1} + p^{i-1}}{2} \right) - B^* p^i, \frac{I_d u^i - u^i}{2} \right) \right| \\ &= R_1 + R_2 + R_3 + R_4. \end{aligned}$$

With the abbreviation $d_i := \alpha u^i + B^* p^i$ we obtain (see, e.g. [3])

$$\begin{aligned} (d_i, I_d u^i - u^i) &= \sum_{K \in \mathcal{T}_h} (d_i, I_d u^i - u^i)_{L^2(K)} \\ &= \sum_{K \in \mathcal{T}_{h,i}^3} (d_i, I_d u^i - u^i)_{L^2(K)}, \end{aligned}$$

since there hold $I_d u^i = u^i$ on $\mathcal{T}_{h,i}^1$ and $d_i = 0$ on $\mathcal{T}_{h,i}^2$ due to $\alpha u^i + B^* p^i = 0$. For every $K \in \mathcal{T}_{h,i}^3$ there is a point x_K with $d_i(x_K) = 0$. Thus, we get (see, e.g. [3])

$$\begin{aligned} |(d_i, I_d u^i - u^i)_{L^2(K)}| &\leq |K|^{1-\frac{2}{s}} \|d_i\|_{L^s(K)} \|I_d u^i - u^i\|_{L^s(K)} \\ &= |K|^{1-\frac{2}{s}} \|d_i - d_i(x_K)\|_{L^s(K)} \|I_d u^i - u^i\|_{L^s(K)} \\ &\leq Ch^2 |K|^{1-\frac{2}{s}} \|\nabla d_i\|_{L^s(K)} \|\nabla u(t_i)\|_{L^s(K)}. \end{aligned}$$

Using the Assumption 3.1, we have

$$\begin{aligned} |(d_i, I_d u^i - u^i)| &\leq Ch^2 \sum_{K \in \mathcal{T}_{h,i}^3} |K|^{1-\frac{2}{s}} \|\nabla d_i\|_{L^s(K)} \|\nabla u(t_i)\|_{L^s(K)} \\ &\leq Ch^2 \left(\sum_{K \in \mathcal{T}_{h,i}^3} |K| \right)^{1-\frac{2}{s}} \|\nabla d_i\|_{L^s(\Omega)} \|\nabla u(t_i)\|_{L^s(\Omega)} \\ &\leq Ch^{3-\frac{2}{s}} \|\nabla d_i\|_{L^s(\Omega)} \|\nabla u(t_i)\|_{L^s(\Omega)}. \end{aligned}$$

We obtain

$$R_1 \leq Ch^{3-\frac{2}{s}} \sum_{i=1}^N \Delta t \|\nabla d_i\|_{L^s(\Omega)} \|\nabla u(t_i)\|_{L^s(\Omega)},$$

$$R_2 \leq Ch^{3-\frac{2}{s}} \sum_{i=1}^N \Delta t \|\nabla d_{i-1}\|_{L^s(\Omega)} \|\nabla u(t_{i-1})\|_{L^s(\Omega)}.$$

For the third term R_3 , we derive that

$$\begin{aligned} R_3 &= C \sum_{i=0}^N \Delta t \left(\alpha \frac{u^{i+1} + u^{i-1}}{2} - \alpha u^i, \frac{I_d u^i - u^i}{2} \right) \\ &\leq C \sum_{i=0}^N \Delta t \left\| \alpha \frac{u^{i+1} + u^{i-1}}{2} - \alpha u^i \right\|_{L^2(\Omega)} \left\| \frac{I_d u^i - u^i}{2} \right\|_{L^2(\Omega)} \\ &\leq C \sum_{i=0}^N \Delta t^2 \alpha \int_{t_i}^{t_{i+1}} \|u_{tt}\|_{L^2(\Omega)} dt \|I_d u^i - u^i\|_{L^2(\Omega)} \\ &\leq C \Delta t^{\frac{3}{2}} \left(\sum_{i=0}^N \left(\int_{t_i}^{t_{i+1}} \|u_{tt}\|_{L^2(\Omega)} dt \right)^2 \right)^{\frac{1}{2}} \left(\sum_{i=0}^N \Delta t \|I_d u^i - u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \\ &\leq C \Delta t^2 \left(\int_0^T \|u_{tt}\|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}} \left(\sum_{i=0}^N \Delta t \|I_d u^i - u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

For the fourth term R_4 , we have

$$R_4 \leq C \Delta t^2 \left(\int_0^T \|p_{tt}\|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}} \left(\sum_{i=0}^N \Delta t \|I_d u^i - u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}.$$

We obtain

$$\begin{aligned} J_3 &\leq Ch^{3-\frac{2}{s}} \|\nabla d\|_{L^2(0,T;L^s(\Omega))} \|\nabla u(t)\|_{L^2(0,T;L^s(\Omega))} \\ &\quad + C \Delta t^2 \left(\int_0^T \|u_{tt}\|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}} \left(\sum_{i=0}^N \Delta t \|I_d u^i - u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \\ &\quad + C \Delta t^2 \left(\int_0^T \|p_{tt}\|_{L^2(\Omega)}^2 dt \right)^{\frac{1}{2}} \left(\sum_{i=0}^N \Delta t \|I_d u^i - u^i\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

For each $m = 0, 1, \dots, N$, we get (see, e.g. [3])

$$\begin{aligned}
\| I_d u(t_m) - u(t_m) \|^2 &= \sum_{K \in \mathcal{T}_h} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} \\
&= \sum_{K \in \mathcal{T}_{h,m}^1} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} \\
&= \sum_{K \in \mathcal{T}_{h,m}^2} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} \\
&\quad + \sum_{K \in \mathcal{T}_{h,m}^3} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)}.
\end{aligned}$$

Since there holds $I_d u(t_m) = u(t_m)$ on $\mathcal{T}_{h,m}^1$, the first term can be estimated as

$$\sum_{K \in \mathcal{T}_{h,m}^1} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} = 0.$$

Since $u(t_m) = -\frac{1}{\alpha} B^* p(t_m)$ on all cells $K \in \mathcal{T}_{h,m}^2$, the second term can be estimated as

$$\begin{aligned}
\sum_{K \in \mathcal{T}_{h,m}^2} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} &\leq C h^4 \sum_{K \in \mathcal{T}_{h,m}^2} \| \nabla^2 u(t_m) \|^2_{L^2(K)} \\
&\leq C h^4 \| \nabla^2 p(t_m) \|^2_{L^2(\Omega)}.
\end{aligned}$$

Using the Hölder inequality and Assumption 3.1, the third term can be estimated as

$$\begin{aligned}
\sum_{K \in \mathcal{T}_{h,m}^3} \| I_d u(t_m) - u(t_m) \|^2_{L^2(K)} &\leq \sum_{K \in \mathcal{T}_{h,m}^3} |K|^{1-\frac{2}{s}} \| I_d u(t_m) - u(t_m) \|^2_{L^s(K)} \\
&\leq C h^2 \sum_{K \in \mathcal{T}_{h,m}^3} |K|^{1-\frac{2}{s}} \| \nabla u(t_m) \|^2_{L^s(K)} \\
&\leq C h^2 \left(\sum_{K \in \mathcal{T}_{h,m}^3} |K| \right)^{1-\frac{2}{s}} \| \nabla u(t_m) \|^2_{L^s(\Omega)} \\
&\leq C h^{3-\frac{2}{s}} \| \nabla u(t_m) \|^2_{L^s(\Omega)}.
\end{aligned}$$

We obtain

$$\sum_{i=1}^N \Delta t \| I_d u(t_i) - u(t_i) \|^2 \leq C h^{3-\frac{2}{s}} \max_{1 \leq i \leq N} \| \nabla u(t_i) \|^2_{L^s(\Omega)} + C h^4 \max_{1 \leq i \leq N} \| \nabla^2 p(t_i) \|^2_{L^2(\Omega)}. \quad (3.23)$$

Combining T_1, T_2 and (3.23) we get

$$\begin{aligned}
\| |u - u_h| \| &\leq C(h^2 + \Delta t^2) + C \Delta t^2 \left(\int_0^T \| p_{tt} \|^2_{L^2(\Omega)} dt \right)^{\frac{1}{2}} \\
&\quad + h^{\frac{3}{2}-\frac{1}{s}} \| \nabla d \|^{\frac{1}{2}}_{L^2(0,T;L^s(\Omega))} \| \nabla u(t) \|^{\frac{1}{2}}_{L^2(0,T;L^s(\Omega))} + C \Delta t^2 \left(\int_0^T \| u_{tt} \|^2_{L^2(\Omega)} dt \right)^{\frac{1}{2}} \\
&\quad + C h^{\frac{3}{2}-\frac{1}{s}} \max_{1 \leq i \leq N} \| \nabla u(t_i) \|^2_{L^s(\Omega)} + C h^2 \max_{1 \leq i \leq N} \| \nabla^2 p(t_i) \|^2_{L^2(\Omega)}. \quad \square
\end{aligned}$$

Theorem 3.9. Let (y, p, u) and (Y_h, P_h, U_h) be the solutions of problems (2.2)-(2.4) and (2.7)-(2.9), respectively. Then we have the following a priori error estimate

$$\| y^n - Y_h^n \|_{L^2(\Omega)} + \| p^n - P_h^n \|_{L^2(\Omega)} \leq C(h^{\frac{3}{2}-\frac{1}{s}} + \Delta t^2), \quad 0 \leq n \leq N.$$

Proof. Let $p_h^n(u)$ and $y_h^n(u)$ be the solutions of (3.6) and (3.5), respectively, satisfying the assumption of Theorem 3.4. Using the triangle inequality, we have

$$\begin{aligned}
\| y^n - Y_h^n \|_{L^2(\Omega)} &\leq \| y^n - y_h^n(u) \|_{L^2(\Omega)} + \| y_h^n(u) - Y_h^n \|_{L^2(\Omega)}, \\
\| p^n - P_h^n \|_{L^2(\Omega)} &\leq \| p^n - p_h^n(u) \|_{L^2(\Omega)} + \| p_h^n(u) - P_h^n \|_{L^2(\Omega)}.
\end{aligned}$$

Using Theorem 3.5, we obtain

$$\begin{aligned}
\| y^n - Y_h^n \|_{L^2(\Omega)} &\leq \| y^n - y_h^n(u) \|_{L^2(\Omega)} + C \| |u - U_h| \|, \\
\| p^n - P_h^n \|_{L^2(\Omega)} &\leq \| p^n - p_h^n(u) \|_{L^2(\Omega)} + C \| |u - U_h| \|.
\end{aligned}$$

Therefore, the result follows from Theorem 3.3 and Theorem 3.8. \square

Table 1
The L^2 error estimates of the control, state, and adjoint states.

N	control		state		adjoint state	
	$\ u - U_h\ $	Order	$\ y - Y_h\ _{L^2(\Omega)}$	Order	$\ p - P_h\ _{L^2(\Omega)}$	Order
4	4.7881×10^{-2}	\	2.0642×10^{-1}	\	7.2354×10^{-2}	\
8	1.4729×10^{-2}	1.7008	5.5176×10^{-2}	1.9035	1.8878×10^{-2}	1.9383
16	4.0332×10^{-3}	1.8686	1.4022×10^{-2}	1.9763	4.7778×10^{-3}	1.9823
32	1.0521×10^{-3}	1.9386	3.5198×10^{-3}	1.9941	1.1982×10^{-3}	1.9954
64	2.6851×10^{-4}	1.9702	8.8308×10^{-4}	1.9948	2.9981×10^{-4}	1.9987

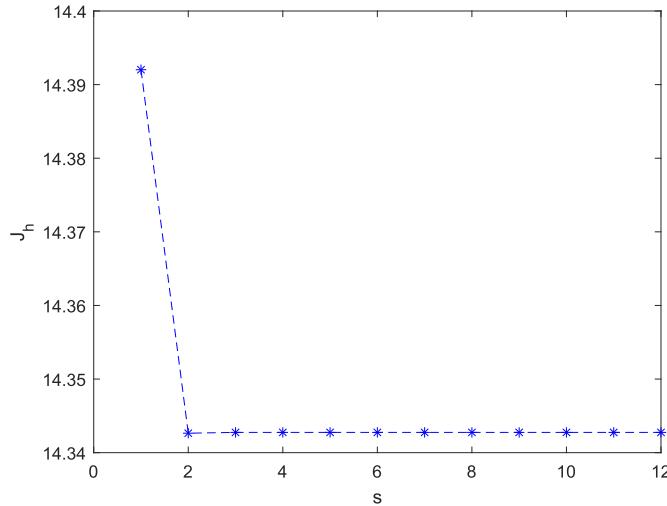


Fig. 1. Decay of the discrete objective functional J_h .

4. Numerical example

In this section we will carry out some numerical experiments to support our theoretical results. We solve the following parabolic optimal control problem:

$$\begin{cases} \min_{u \in U_{ad}} \frac{1}{2} \|y - y_d\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{1}{2} \|u\|_{L^2(0,T;L^2(\Omega))}^2 \\ y_t - \Delta y = f + u, \quad (x,t) \in \Omega \times (0,T], \\ y(\cdot, 0) = y_0, \quad x \in \Omega. \end{cases}$$

Here

$$U_{ad} = \{u \in L^2(0,T;L^2(\Omega)), u_a \leq u(t,x) \leq u_b, \text{ a.e. in } \Omega, t \in (0,T], u_a, u_b \in \mathbb{R}\}.$$

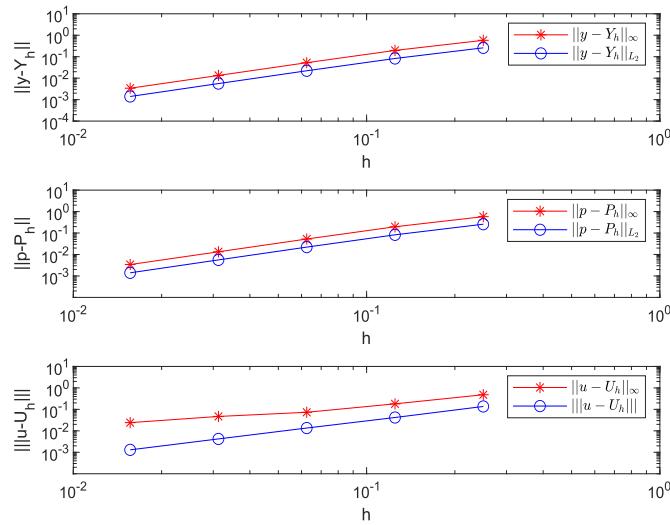
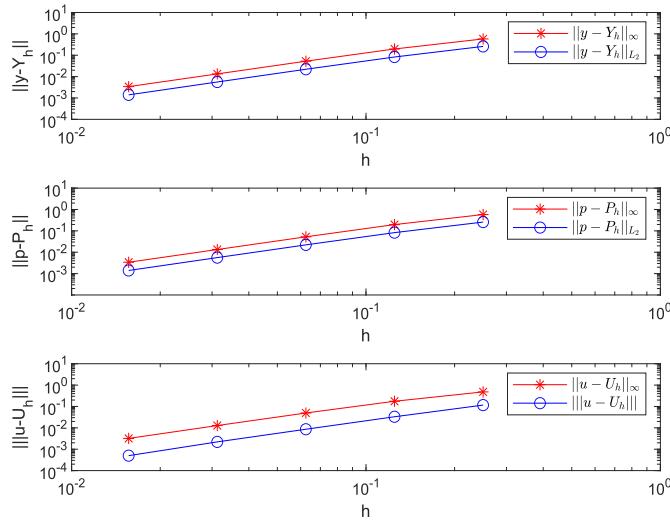
Example 4.1. Let Ω be a square domain $(0,1) \times (0,1)$, $T = 1$. The data are as follows:

$$\begin{aligned} u_a &= -1, u_b = 1, \\ y(x,t) &= \exp(t) \sin(\pi x) \sin(\pi y), \\ u(x,t) &= \max\{u_a, \min\{u_b, (T-t)^2 \sin(\pi x) \sin(\pi y)\}\}, \\ p(x,t) &= -(T-t)^2 \sin(\pi x) \sin(\pi y), \\ f(x,t) &= y_t(x,t) - \Delta y(x,t) - u(x,t), \\ y_d(x,t) &= y(x,t) + p_t(x,t) + \Delta p(x,t). \end{aligned}$$

The errors on a sequence of uniformly refined meshes are shown in Table 1. Fig. 1 shows the variation in the value of the objective function J_h with the number of iterations s in the case for $h = 1/16$ and $\Delta t = 1/16$. The figure shows that the value of the functional J_h converges to its minimum as the number of iterations increases.

Example 4.2. Let Ω be the square domain $(0,1) \times (0,1)$, $T = 1$. The data were as follows:

$$\begin{aligned} u_a &= -0.4, u_b = 0.4, \\ y(x,t) &= t \sin(2\pi x) \sin(2\pi y), \\ u(x,t) &= \max\{u_a, \min\{u_b, -(T-t) \sin(2\pi x) \sin(2\pi y)\}\}, \end{aligned}$$

Fig. 2. Discretization error y, u, p .Fig. 3. Discretization error y, u, p .

$$p(x, t) = (T - t) \sin(2\pi x) \sin(2\pi y),$$

$$f(x, t) = y_t(x, t) - \Delta y(x, t) - u(x, t),$$

$$y_d(x, t) = y(x, t) + p_t(x, t) + \Delta p(x, t).$$

We show the convergence rates of the control, state, and adjoint in Fig. 2 with piecewise linear discretization, and in Fig. 3 with variational discretization. In Fig. 4, we plot the profile of the numerical solution, U_h at $t = 0.5$ when $\Delta t = 1/64$ and $h = 1/64$.

Example 4.3. Let Ω be a cube domain $(0, 1) \times (0, 1) \times (0, 1)$, $T = 1$. The data are as follows

$$u_a = -1, u_b = 1,$$

$$y(x, t) = \exp(t) \sin(\pi x) \sin(\pi y) \sin(\pi z),$$

$$u(x, t) = \max\{u_a, \min\{u_b, (T - t)^2 \sin(\pi x) \sin(\pi y) \sin(\pi z)\}\},$$

$$p(x, t) = -(T - t)^2 \sin(\pi x) \sin(\pi y) \sin(\pi z),$$

$$f(x, t) = y_t(x, t) - \Delta y(x, t) - u(x, t),$$

$$y_d(x, t) = y(x, t) + p_t(x, t) + \Delta p(x, t).$$

Table 2 shows the errors and convergence order of state variable y , dual state variable p and control variable u .

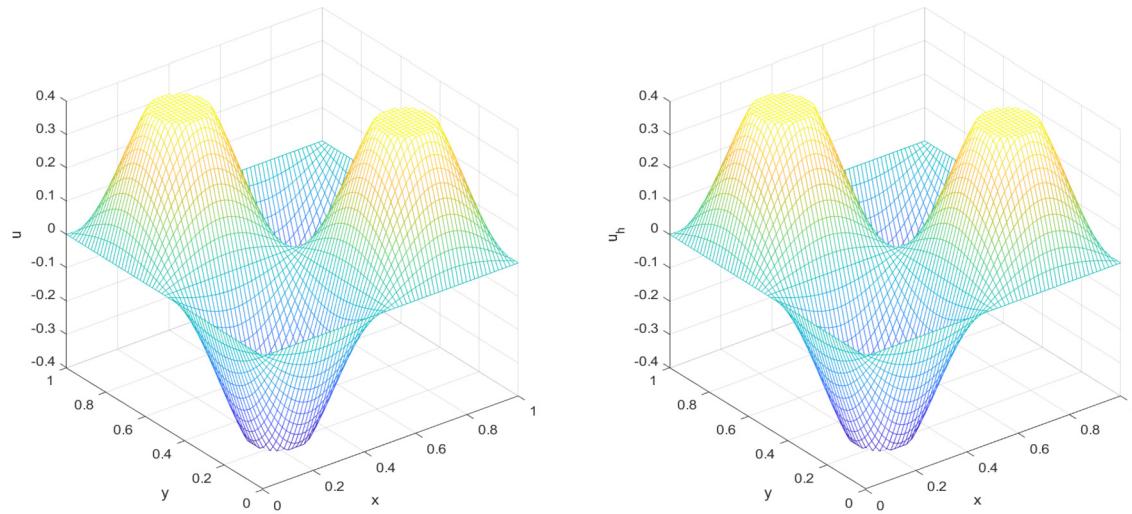


Fig. 4. The exact solution u (left) and the numerical solution U_h (right) at $t = 0.5$.

Table 2
The L^2 error estimates of the control, state, and adjoint states.

N	control		state		adjoint state	
	$\ u - U_h\ $	Order	$\ y - Y_h\ _{L^2(\Omega)}$	Order	$\ p - P_h\ _{L^2(\Omega)}$	Order
2	1.3101×10^{-1}	\	6.3529×10^{-1}	\	2.3555×10^{-1}	\
4	8.0177×10^{-2}	0.7085	2.8584×10^{-1}	1.1522	1.0483×10^{-1}	1.1680
8	2.7952×10^{-2}	1.5202	8.6348×10^{-2}	1.7270	3.1728×10^{-2}	1.7242
16	7.9007×10^{-3}	1.8230	2.2844×10^{-2}	1.9184	8.3954×10^{-3}	1.9181

5. Conclusion

In this paper, we discuss the Crank-Nicolson finite element method for parabolic optimal control problems with pointwise inequality constraints on the control variable. The state and adjoint state are approximated using piecewise linear functions. For control discretization, we considered piecewise linear finite element discretization or variational discretization. The corresponding priori error estimates were obtained. Numerical experiments verified the theoretical analysis. Furthermore, many practical remain problems to be solved, such as nonlinear optimal control and boundary control.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was partially supported by grants from the Xinjiang Uygur Autonomous Region Natural Science Foundation (No. 2022D01C409), the National Natural Science Foundation of China (No. 61962056), and the Xinjiang Uygur Autonomous Region Natural Science Fund (No. 2019D01C047).

References

- [1] I. Lasiecka, K. Malanowski, On discrete-time Ritz-Galerkin approximation of control constrained optimal control problems for parabolic equations, *Control Cybern.* 7 (1) (1978) 21–36.
- [2] D. Meidner, B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems part I: problems without control constraints, *SIAM J. Control Optim.* 47 (3) (2008) 1150–1177.
- [3] D. Meidner, B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems part II: problems with control constraints, *SIAM J. Control Optim.* 47 (3) (2008) 1301–1329.
- [4] Y.L. Tang, Y.P. Chen, Superconvergence of finite element methods for optimal control problems governed by parabolic equations with time-dependent coefficients, *East Asian J. Appl. Math.* 3 (3) (2013) 209–227.
- [5] Y.P. Chen, Superconvergence of mixed finite element methods for optimal control problems, *Math. Comput.* 77 (263) (2008) 1269–1291.
- [6] Y.P. Chen, Y.Q. Huang, N.Y. Yi, A posteriori error estimates of spectral method for optimal control problems governed by parabolic equations, *Sci. China Ser. A, Math.* 55 (8) (2008) 1376–1390.
- [7] R.S. Falk, Approximation of a class of optimal control problems with order of convergence estimates, *J. Math. Anal. Appl.* 44 (1) (1973) 28–47.
- [8] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case, *Comput. Optim. Appl.* 30 (1) (2005) 45–61.
- [9] Y.P. Chen, Y.Q. Huang, W.B. Liu, N.N. Yan, Error estimates and superconvergence of mixed finite element methods for convex optimal control problems, *J. Sci. Comput.* 42 (3) (2010) 382–403.
- [10] Z.L. Lu, Y.P. Chen, W.S. Zheng, A posteriori error estimates of lowest order Raviart-Thomas mixed finite element methods for bilinear optimal control problems, *East Asian J. Appl. Math.* 2 (2) (2012) 108–125.

- [11] J.W. Zhou, Y.P. Chen, Y.Q. Dai, Superconvergence of triangular mixed finite elements for optimal control problems with an integral constraint, *Appl. Math. Comput.* 217 (5) (2010) 2057–2066.
- [12] Y.P. Chen, Y.Q. Dai, Superconvergence for optimal control problems governed by semi-linear elliptic equations, *J. Sci. Comput.* 39 (2) (2009) 206–221.
- [13] Y.P. Chen, W.B. Liu, Error estimates and superconvergence of mixed finite element for quadratic optimal control, *Int. J. Numer. Anal. Model.* 3 (3) (2006) 311–321.
- [14] Y.P. Chen, T.L. Hou, Error estimates and superconvergence of RT0 mixed methods for a class of semilinear elliptic optimal control problems, *Numer. Math., Theory Methods Appl.* 6 (4) (2013) 637–656.
- [15] L. Bonifacius, K. Pieper, B. Vexler, A priori error estimates for space-time finite element discretization of parabolic time-optimal control problems, *SIAM J. Control Optim.* 57 (1) (2019) 129–162.
- [16] X.B. Luo, Y.P. Chen, Y.G. Huang, A priori error estimates of Crank-Nicolson finite volume element method for parabolic optimal control problems, *Adv. Appl. Math. Mech.* 5 (5) (2013) 688–704.
- [17] Y. Tang, Y. Hua, Superconvergence of fully discrete finite elements for parabolic control problems with integral constraints, *East Asian J. Appl. Math.* 3 (2) (2013) 138–153.
- [18] C. Meyer, A. Rösch, Superconvergence properties of optimal control problems, *SIAM J. Control Optim.* 43 (3) (2004) 970–985.
- [19] Y.L. Tang, Y.P. Chen, Superconvergence analysis of fully discrete finite element methods for semilinear parabolic optimal control problems, *Front. Math. China* 8 (2) (2013) 443–464.
- [20] Y.L. Tang, Y.P. Chen, Recovery type a posteriori error estimates of fully discrete finite element methods for general convex parabolic optimal control problems, *Numer. Math., Theory Methods Appl.* 5 (4) (2012) 573–591.
- [21] W. Gong, M. Hinze, Z.J. Zhou, Finite element method and a priori error estimates for Dirichlet boundary control problems governed by parabolic PDEs, *J. Sci. Comput.* 66 (3) (2016) 941–967.
- [22] D. Meidner, B. Vexler, A priori error analysis of the Petrov-Galerkin Crank-Nicolson scheme for parabolic optimal control problems, *SIAM J. Control Optim.* 49 (5) (2011) 2183–2211.
- [23] T. Apel, T.G. Flaig, Crank-Nicolson schemes for optimal control problems with evolution equations, *SIAM J. Numer. Anal.* 50 (3) (2012) 1484–1512.
- [24] A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Publishing Company, North-Holland, 1994.
- [25] J.G. Heywood, R. Rannacher, Finite-element approximation of the nonstationary Navier-Stokes problem part IV: error analysis for second-order time discretization, *SIAM J. Numer. Anal.* 27 (2) (1990) 353–384.
- [26] N.V. Daniels, M. Hinze, M. Vierling, Crank-Nicolson time stepping and variational discretization of control-constrained parabolic optimal control problems, *SIAM J. Control Optim.* 53 (3) (2015) 1182–1198.