Finite Element Methods for An Optimal Control Problem Governed by Heat Equation with Bilateral Controls

Ruirui Zheng, Keying Ma*

Abstract—**The utilization of finite element technique is explored for an optimization problem by heat conduction equation with bilateral controls, which are the distributed and boundary control, respectively. The derivation of the co-state equation and optimality conditions is accomplished through the application of optimal control theory. To set up the fully discrete approximation schemes, piecewise linear continuous functions are employed for the approximation of state and co-state variables, whereas piecewise constant for control variable approximation. A priori error estimates are rigorously established for all considered variables under appropriate norms. Theoretical findings are validated by the presentation of comprehensive numerical experiments.**

Index Terms—**optimal control problem; heat equation; co-state variable; optimality conditions; priori error estimates.**

I. Introduction

THE significant importance of optimal control problems (OCPs) constrained by partial differential equations (PDEs) is widely acknowledged across various physical and engineering applications. Solving PDEs-constrained OCPs is fundamentally equivalent to determining the solution of an optimality system, which typically comprises three essential components involving the state, co-state, and optimality formulations. For the systematic introduction of these problems, readers are referred to several fundamental references, including [1]–[3].

Finite element method (FEM) has been extensively employed for discussing PDE-constrained optimization problems, primarily through discrete approximation to the corresponding optimality system. Considerable academic attention has increasingly focused on this topic recently, yielding numerous valuable contributions. Among these, space-time FEM was developed for parabolic type OCPs, with and without control constrains in [4] and [5], respectively. Furthermore, reference [6] presented a characteristic FEM for OCPs, which were subject to transient advection-diffusion equations. Other numerical approaches were also proposed, including mixed FEM [7]–[9], semi-discrete FEM [10], and discontinuous Galerkin FEM [11]. The aforementioned studies predominantly concentrated on establishing a priori error analysis under appropriate norms. Also, substantial progress was achieved in the investigation of a posteriori error estimates of using FEM for PDE-constrained OCPs. For this aspect, readers are referred to references [2], [12]–[17] and the closely related literatures cited by them.

Mixed optimal controls arose from optimization problems in efficient building operation, for examples, [18]- [19]. Despite their practical significance, research on mixed optimal controls using FEM remains relatively unexplored. In [20], a mixed-integer OCPs constrained by heat equation was investigated, involving both continuous and discrete controls. To tackle these mixed optimal controls, a modified branch-and-bound method [21]- [22] was employed, and POD method [23]- [26] was also accomplished to reduce the model order. However, it should be noted that [20] did not discuss a priori error estimates of FEM. Here, we will consider a heat equation constrained OCP with mixed controls, similar to [20], but with a distinct feature: the controls are applied both in the interior and on the boundary of the domain. Our aim is to attain a desired temperature distribution with minimal energy consumption by optimizing the domain and boundary controls, which represent the internal heating strategies and boundary insulation materials, respectively. Unlike [20], the boundary controls in our study are continuous. We primarily focus on developing the fully discrete approximation schemes and rigorously deriving a priori error estimates through detailed theoretical analysis.

The structure of the subsequent sections is as follows. Section II describes the formulation of an optimization model in detail, then derive the optimality system through the optimal control theory. In Section III, we set up the fully discrete schemes. Section IV contributes to derive a priori error estimates for all considered variables by using suitable norms. In Section V, numerical experiments for three distinct scenarios are carefully performed and comprehensively analyzed to confirm our theoretical findings. In the end, Section VI provides some concluding remarks.

II. An optimization model

Our subsequent investigation will concentrate on a heat conduction equation:

$$\begin{cases} y_t(t,x) - \mu\Delta y(t,x) \\ \quad = \sum_{j=1}^{N^c} B_j u_j^c(t,x),\ (t,x)\in Q, \\ \kappa\dfrac{\partial y}{\partial n}(t,x) = \displaystyle\sum_{j=1}^{N^i} u_j^i(t,x), (t,x)\in\Sigma, \\ y(0,x) = y_0(x),\ x\in\Omega. \end{cases} \tag{1}$$

Manuscript received October 6, 2024; revised March 5, 2025.

This work was supported in part by the Natural Science Foundation of China (Grant No. 11871312), and the Natural Science Foundation of Shandong Province, China (Grant No. ZR2023MA086).

Ruirui Zheng is a lecturer of School of Science, Shandong Jiaotong University, Jinan 250357, China (e-mail:zhengruirui@sdjtu.edu.cn).

*Keying Ma is an associate professor of School of Mathematics, Shandong University, Jinan 250100, China (corresponding author to provide phone: +86-13695416988; fax: +86-531-88364652; e-mail: makeying@sdu.edu.cn).

This equation models the temperature distribution $y(t,x)$ within a building at position $x \in \Omega$ and time $t \in J$, where $\Omega \subset R^d\,(d=2,3)$ is a polygonal domain, $J=[0,T]$ is a time interval, and $T>0$ is the final time. Here, $Q = J\times\Omega$ represents the space-time domain, $\Sigma = J\times\Gamma$ denotes the boundary over time interval, and $\mu \in \mathbb{R}_+ := \{s\in\mathbb{R}\,|\,s>0\}$ is a constant thermal diffusion coefficient. The initial temperature distribution $y_0(x)$ is considered as a known function, belonging to $L^\infty(\Omega)$.

Here, we consider that domain Ω is partitioned into disjoint sub-domains such that $\Omega = \cup_{j=1}^{N^c}\Omega_j$, with the corresponding boundary $\Gamma = \cup_{j=0}^{N^i}\Gamma_j$. Each sub-domain $\Omega_j\,(j\;=1,\cdots,N^c)$ represents an independently controllable heating zone within the building, while each boundary segment $\Gamma_j\,(j=1,\cdots,N^i)$ corresponds to an exterior wall section requiring insulation. For each $j=1,\cdots,N^c$, $u_j^c \in L^2(Q)$ is a control function with support restricted to $Q_j := J\times\Omega_j$, and $B_j:\ L^2(\Omega_j)\to L^2(\Omega)$ is a linear continuous operator. Similarly, for each $j=1,\cdots,N^i$, $u_j^i \in L^2(\Sigma)$ is a boundary control function with support confined to $\Sigma_j := J\times\Gamma_j$. The interior walls Γ_0 is subject to homogeneous Neumann boundary condition, modeling perfect insulation.

The state space is defined as $W = H^1(0,T;L^2(\Omega))\cap L^2(0,T;V)$, where $V=H^1(\Omega)$. Two control spaces are specified as $X^c = L^2(0,T;L^2(\Omega))$ for domain controls and $X^i = L^2(0,T;L^2(\Gamma))$ for boundary controls, respectively. The corresponding constrained admissible controls are two closed convex sets $K^c\subset X^c$ and $K^i\subset X^i$. Specifically, the control vectors $u^c=\{u_j^c\}_{j=1}^{N^c}\in K^c$ and $u^i=\{u_j^i\}_{j=1}^{N^i}\in K^i$ represent heating strategies over Q and insulation strategies on Σ, respectively.

Based on the framework introduced in [20], we propose the following OCP subject to (1): Find $(u^c,\,u^i)\in K^c\times K^i$ such that

$$\begin{aligned}\min_{u^c\in K^c,\,u^i\in K^i}&\Big\{J(u^c,\,u^i)\\&=\frac{\alpha^Q}{2}\int_Q|y(t,x)-y_d(t,x)|^2dxdt\\&+\frac{1}{2}\sum_{j=1}^{N^c}\alpha_j^c\int_{Q_j}|u_j^c(t,x)|^2dxdt\\&+\frac{1}{2}\sum_{j=1}^{N^i}\alpha_j^i\int_{\Sigma_j}|u_j^i(t,s)-\hat{u}_j^i(t,s)|^2dsdt\Big\}.\end{aligned}\tag{2}$$

The positive parameters $\alpha^c=\{\alpha_j^c\}_{j=1}^{N^c}$ and $\alpha^i=\{\alpha_j^i\}_{j=1}^{N^i}$ are the cost weights associated with domain and boundary controls, respectively. We suppose the existence of two positive constants α_*^c and α_*^i satisfying $\min\limits_{1\le j\le N^c}\alpha_j^c>\alpha_*^c$ and $\min\limits_{1\le j\le N^i}\alpha_j^i>\alpha_*^i$. Here, $\hat{u}^i=\{\hat{u}_j^i\}_{j=1}^{N^i}$ denotes the minimal-cost insulation configuration. The target temperature distribution is given by $y_d\in L^2(Q)$ and the positive constant α^Q quantifies the relative importance of temperature regulation in the cost functional. The objective of the functional (2) is designed to approximate the desired temperature distribution y_d as closely as possible and simultaneously minimize the heating and insulation costs through optimal selection of domain controls u^c and boundary controls u^i.

Let

$$\begin{aligned}(\mu\nabla v,\nabla w)&=\int_\Omega\mu\nabla v\nabla w dx,\;\forall\, v,\,w\in H^1(\Omega),\\(\phi_1,\phi_2)&=\int_\Omega\phi_1\phi_2dx,\;\forall\,\phi_1,\,\phi_2\in L^2(\Omega),\\(\psi_1,\psi_2)_{\Gamma_j}&=\int_{\Gamma_j}\psi_1\psi_2ds,\;\forall\,\psi_1,\,\psi_2\in L^2(\Gamma_j),\\&\qquad j=1,\cdots,N^i.\end{aligned}$$

As demonstrated in [1]- [2], problem (1)-(2) admits a unique solution triplet $\{y,u^c,u^i\}\in W\times K^c\times K^i$ if and only if a co-state variable $p\in W$ exists. In this case, the quadruple $\{y,p,u^c,u^i\}\in W\times W\times K^c\times K^i$ satisfies the optimality system below, referred to as **(OCP-OPT)**:

$$\begin{cases}(y_t,\omega)+(\mu\nabla y,\nabla\omega)\\\quad=\displaystyle\sum_{j=1}^{N^c}\big(B_ju_j^c,\omega\big)+\sum_{j=1}^{N^i}\big(u_j^i,\omega\big)_{\Gamma_j},\;\forall\;\omega\in V,\\y(0,x)=y_0(x),\end{cases}\tag{3}$$

$$\begin{cases}(-p_t,q)+(\mu\nabla q,\nabla p)\\\qquad=\big(\alpha^Q(y-y_d),q\big),\;\forall\;q\in V,\\p(T,x)=0,\end{cases}\tag{4}$$

$$\begin{aligned}&\int_0^T(\alpha_j^cu_j^c+B_j^*p,v_j^c-u_j^c)dt\ge0,\\&\qquad j=1,\cdots,N^c,\quad\forall\;v^c\in K^c,\end{aligned}\tag{5}$$

$$\begin{aligned}&\int_0^T\big(\alpha_j^i(u_j^i-\hat{u}_j^i)+p,v_j^i-u_j^i\big)_{\Gamma_j}dt\ge0,\\&\qquad j=1,\cdots,N^i,\quad\forall\;v^i\in K^i,\end{aligned}\tag{6}$$

where B_j^* is the adjoint operator of B_j. The inequalities (5)-(6) are the optimality conditions.

III. Fully discrete approximate schemes

We begin to address the fully discrete schemes for the system **(OCP-OPT)**. We divide $[0,T]$ into subintervals defined by $0=t^0<t^1<\cdots<t^{N_T}=T$, where $\Delta t^k=t^k-t^{k-1}$ for $k=1,\cdots,N_T$, and $\Delta t=\max\limits_{1\le k\le N_T}\{\Delta t^k\}$. For integers $0\le s\le1$ and $1\le p<\infty$, we introduce the discrete time-dependent norms as follows:

$$\|f\|_{l^p(J;H^s(\Omega))}=\Big(\sum_{k=1}^{N_T}\Delta t^k\|f^k\|_{H^s(\Omega)}^p\Big)^{\frac{1}{p}},$$

$$\|f\|_{l^\infty(J;H^s(\Omega))}=\max_{0\le k\le N_T}\|f^k\|_{H^s(\Omega)}.$$

The norms $\|f\|_{l^p(J;L^2(\Gamma))}$ is defined analogously.

We suppose that the domain Ω has a quasi-uniform triangulation denoted as T_h, satisfying $\Omega=\bigcup_{\tau\in T_h}\overline{\tau}$. Let $h=\max_{\tau\in T_h}h_\tau$, where h_τ is the diameter of the element τ. Associated with T_h, we introduce a finite element space $S_h\subset C(\Omega)$, where for every $\chi\in S_h$ and $\tau\in T_h$, the restriction $\chi|_\tau$ is a first-order polynomial. Based on this spatial discretization, we define two discrete function spaces: $W_h=L^2(J;V_h)$, $V_h=S_h\cap V$.

The discretization can be established for the control spaces similarly. Assume that T_h^c be a quasi-uniform triangulation of Ω, such that $\Omega=\bigcup_{\tau^c\in T_h^c}\overline{\tau}^c$. Let $h_{u^c}=\max_{\tau^c\in T_h^c}h_{\tau^c}$, where h_{τ^c} represents the diameter of the element τ^c. Corresponding to T_h^c, another finite element space $U_h^c \subset L^2(\Omega)$ is defined such that for every $\chi \in U_h^c$ and $\tau^c \in T_h^c$, the restriction $\chi|_{\tau^c}$ is a zero-order polynomial. Let $K_h^c = U_h^c \cap K^c$. It is easy to see that $K_h^c \subset K^c$. Let T_h^i be a quasi-uniform triangulation of Γ satisfying $\Gamma = \bigcup_{\tau^i \in T_h^i} \overline{\tau}^i$. Let $h_{u^i} = \max_{\tau^i \in T_h^i} h_{\tau^i}$, where h_{τ^i} is the diameter of the element τ^i. Corresponding to T_h^i, a finite element space $U_h^i \subset L^2(\Gamma)$ is constructed such that $\chi|_{\tau^i}$ is a zero-order polynomial for every $\chi \in U_h^i$ and $\tau^i \in T_h^i$. Let $K_h^i = U_h^i \cap K^i$. Clearly, K_h^i is a subset of K^i.

The fully discrete approximate schemes for (1)-(2) are to find $\{Y^k, U^{ck}, U^{ik}\} \in V_h \times K_h^c \times K_h^i, k = 1, \cdots, N_T$, such that

$$\begin{aligned}\min_{U^{ck} \in K_h^c,\, U^{ik} \in K_h^i} &\Big\{ J_h\big(U^{ck}, U^{ik}\big) \\ &= \frac{\alpha^Q}{2} \sum_{k=1}^{N_T} \|Y^k - y_d^k\|^2_{L^2(\Omega)} \Delta t^k \\ &+ \frac{\alpha_j^c}{2} \sum_{k=1}^{N_T} \sum_{j=1}^{N^c} \|U_j^{ck}\|^2_{L^2(\Omega_j)} \Delta t^k \\ &+ \frac{\alpha_j^i}{2} \sum_{k=1}^{N_T} \sum_{j=1}^{N^i} \|U_j^{ik} - \hat{u}_j^{ik}\|^2_{L^2(\Gamma_j)} \Delta t^k \Big\},\end{aligned} \tag{7}$$

subject to

$$\begin{cases} \Big(\dfrac{Y^k - Y^{k-1}}{\Delta t^k}, \omega_h\Big) + (\mu \nabla Y^k, \nabla \omega_h) \\ = \displaystyle\sum_{j=1}^{N^c} (B_j U_j^{ck}, \omega_h) + \sum_{j=1}^{N^i} (U_j^{ik}, \omega_h)_{\Gamma_j}, \forall\, \omega_h \in V^h, \\ Y^0 = y_h^0, \end{cases} \tag{8}$$

where $y_h^0 \in V^h$ is an approximation to $y_0(x)$.

By the optimal control theory [1]- [2], for $k = 1, \cdots, N_T$, problem (7)-(8) has a unique solution triplet $\{Y^k, U^{ck}, U^{ik}\} \in V_h \times K_h^c \times K_h^i$, if and only if a co-state $P^{k-1} \in V_h$ exists. In this case, the quadruple $\{Y^k, P^{k-1}, U^{ck}, U^{ik}\} \in V_h \times V_h \times K_h^c \times K_h^i$ satisfies the following discretized system, denoted as **(OCP-OPT)**^{h}

$$\begin{cases} \Big(\dfrac{Y^k - Y^{k-1}}{\Delta t^k}, \omega_h\Big) + (\mu \nabla Y^k, \nabla \omega_h) \\ = \displaystyle\sum_{j=1}^{N^c} (B_j U_j^{ck}, \omega_h) + \sum_{j=1}^{N^i} (U_j^{ik}, \omega_h)_{\Gamma_j}, \forall\, \omega_h \in V^h, \\ Y^0 = y_h^0. \end{cases} \tag{9}$$

$$\begin{cases} \Big(\dfrac{P^{k-1} - P^k}{\Delta t^k}, q_h\Big) + (\mu \nabla q_h, \nabla P^{k-1}) \\ \qquad = \big(\alpha^Q (Y^k - y_d^k), q_h\big), \qquad \forall\, q_h \in V^h, \\ P^{N_T} = 0, \end{cases} \tag{10}$$

$$\Big(\alpha_j^c U_j^{ck} + B_j^* P^{k-1}, v_j^c - U_j^{ck}\Big) \geq 0, \qquad j = 1, \cdots, N^c, \quad \forall\, v_j^c \in K_h^c, \tag{11}$$

$$\Big(\alpha_j^i (U_j^{ik} - \hat{u}_j^{ik}) + P^{k-1}, v_j^i - U_j^{ik}\Big)_{\Gamma_j} \geq 0, \qquad j = 1, \cdots, N^i, \quad \forall\, v_j^i \in K_h^i. \tag{12}$$

IV. Error analysis

In the process of deriving a priori error estimates for system **(OCP-OPT)**^{h}, we need to use two auxiliary variables $\{Y^k(u), P^k(u)\} \in V_h \times V_h$, $k = 1, \cdots, N_T$, defined as follows:

$$\begin{cases} \Big(\dfrac{Y^k(u) - Y^{k-1}(u)}{\Delta t^k}, \omega_h\Big) + \big(\mu \nabla Y^k(u), \nabla \omega_h\big) \\ = \displaystyle\sum_{j=1}^{N^c} (B_j u_j^{ck}, \omega_h) + \sum_{j=1}^{N^i} (u_j^{ik}, \omega_h)_{\Gamma_j}, \forall\, \omega_h \in V^h, \\ Y^0(u) = y_h^0, \end{cases} \tag{13}$$

$$\begin{cases} \Big(\dfrac{P^{k-1}(u) - P^k(u)}{\Delta t^k}, q_h\Big) + (\mu \nabla q_h, \nabla P^{k-1}(u)) \\ = \Big(\alpha^Q (Y^k(u) - y_d^k), q_h\Big), \;\; \forall\, q_h \in V^h. \\ P^{N_T}(u) = 0. \end{cases} \tag{14}$$

Here, $Y^k(u)$ and $P^k(u)$ can be regarded as the finite element solutions to equations (3) and (4), depending on the exact controls $u^c = \{u_j^c\}_{j=1}^{N^c}$ and $u^i = \{u_j^i\}_{j=1}^{N^i}$. The auxiliary equations (13) and (14) are introduced solely for the purposes of theoretical analysis and commonly employed as a standard technique in the literature, such as [6]- [11].

For simplicity of presentation, the following notations are adopted in this paper:

$$\theta^k = Y^k - Y^k(u), \quad \eta^k = y^k - Y^k(u),$$
$$\zeta^k = P^k - P^k(u), \quad \xi^k = p^k - P^k(u),$$

where $k = 0, 1, \cdots, N_T$. Clearly, we know $\theta^0 = 0$ and $\zeta^{N_T} = 0$. With no loss of generality, we take two positive integers L and M satisfying

$$\|v^L\| := \|v\|_{l^\infty(J; L^2(\Omega))}, \qquad \text{for } v = \theta \text{ and } \eta,$$
$$\|\omega^M\| := \|\omega\|_{l^\infty(J; L^2(\Omega))}, \qquad \text{for } \omega = \zeta \text{ and } \xi.$$

Moreover, positive constant C will represent a general one, whose value is irrelevant to h and Δt, and may vary in different circumstances.

A. Estimates for $Y - Y(u)$ and $P - P(u)$

Lemma 1 Let $\{Y, P\}$ and $\{Y(u), P(u)\}$ be the solutions of (9)-(12) and (13)-(14), respectively. Then, the error estimates below hold:

$$\begin{aligned}&\|Y - Y(u)\|_{l^\infty(J; L^2(\Omega))} + \|\nabla(Y - Y(u))\|_{l^2(J; L^2(\Omega))} \\ &\leq C\big\{\|U^c - u^c\|_{l^2(J; L^2(\Omega))} + \|U^i - u^i\|_{l^2(J; L^2(\Gamma))}\big\},\end{aligned} \tag{15}$$

$$\begin{aligned}&\|P - P(u)\|_{l^\infty(J; L^2(\Omega))} + \|\nabla(P - P(u))\|_{l^2(J; L^2(\Omega))} \\ &\leq C\big\{\|U^c - u^c\|_{l^2(J; L^2(\Omega))} + \|U^i - u^i\|_{l^2(J; L^2(\Gamma))}\big\},\end{aligned} \tag{16}$$

where the following norms are used

$$\begin{cases} \|U^c - u^c\|_{l^2(J; L^2(\Omega))} = \big(\sum\limits_{k=1}^{N_T} \Delta t^k \|U^{ck} - u^{ck}\|^2\big)^{\frac{1}{2}}, \\ \|U^i - u^i\|_{l^2(J; L^2(\Gamma))} = \big(\sum\limits_{k=1}^{N_T} \Delta t^k \|U^{ik} - u^{ik}\|^2_{L^2(\Gamma)}\big)^{\frac{1}{2}}, \end{cases}$$

and

$$\begin{cases} \|U^{ck}-u^{ck}\|^2 := \sum_{j=1}^{N^c} \|U_j^{ck}-u_j^{ck}\|^2_{L^2(\Omega_j)}, \\ \|U^{ik}-u^{ik}\|^2_{L^2(\Gamma)} := \sum_{j=1}^{N^i} \|U_j^{ik}-u_j^{ik}\|^2_{L^2(\Gamma_j)}. \end{cases}$$

Proof. The derivations of (15) and (16) are similar. Thus, we only provide the proof for (15) here. Subtract (13) from (9) to obtain

$$\Big(\frac{\theta^k-\theta^{k-1}}{\Delta t^k},\omega_h\Big)+(\mu\nabla\theta^k,\nabla\omega_h) = \sum_{j=1}^{N^c}\Big(B_j(U_j^{ck}-u_j^{ck}),\omega_h\Big)+\sum_{j=1}^{N^i}\Big((U_j^{ik}-u_j^{ik}),\omega_h\Big)_{\Gamma_j}. \tag{17}$$

By choosing $\omega_h=\theta^k$ in (17), we obtain

$$\frac{1}{2\Delta t^k}\Big\{\|\theta^k\|^2-\|\theta^{k-1}\|^2+\|\theta^k-\theta^{k-1}\|^2\Big\} +\mu\|\nabla\theta^k\|^2 \le I_1+I_2, \tag{18}$$

and

$$I_1 \le C\sum_{j=1}^{N^c}\|U_j^{ck}-u_j^{ck}\|^2_{L^2(\Omega_j)}+\frac{\varepsilon}{4}\|\theta^k\|^2 \le C\|U^{ck}-u^{ck}\|^2+\frac{\varepsilon}{4}\|\theta^k\|^2, \tag{19}$$

$$I_2 \le C\sum_{j=1}^{N^i}\|U_j^{ik}-u_j^{ik}\|^2_{L^2(\Gamma_j)}+\frac{\mu}{2}\|\nabla\theta^k\|^2+\frac{\varepsilon}{4}\|\theta^k\|^2 \le C\|U^{ik}-u^{ik}\|^2_{L^2(\Gamma)}+\frac{\mu}{2}\|\nabla\theta^k\|^2+\frac{\varepsilon}{4}\|\theta^k\|^2, \tag{20}$$

where the trace theorem is used and positive constant ε can be selected sufficiently small.

Let inequality (18) be multiplied by $2\Delta t^k$ on both sides and summed over k from 1 to L. Performing straightforward calculations and using the condition $\theta^0=0$, we have

$$(1-\varepsilon\Delta t^k)\|\theta^L\|^2+\sum_{k=1}^{L}\|\theta^k-\theta^{k-1}\|^2+\mu\sum_{k=1}^{L}\Delta t^k\|\nabla\theta^k\|^2 \le \varepsilon\sum_{k=1}^{L-1}\Delta t^k\|\theta^k\|^2+C\sum_{k=1}^{L}\Delta t^k\|U^{ck}-u^{ck}\|^2 + C\sum_{k=1}^{L}\Delta t^k\|U^{ik}-u^{ik}\|^2_{L^2(\Gamma)}. \tag{21}$$

Now, we select the previously used ε sufficiently small such that $0<\varepsilon\Delta t^k\le\varepsilon\Delta t\le\frac{1}{2}$ and $1-\varepsilon\Delta t^k\ge\frac{1}{2}$. Hence, we obtain

$$\|\theta^L\|^2+2\sum_{k=1}^{L}\|\theta^k-\theta^{k-1}\|^2+2\mu\sum_{k=1}^{L}\Delta t^k\|\nabla\theta^k\|^2 \le C\sum_{k=1}^{L-1}\Delta t^k\|\theta^k\|^2+C\sum_{k=1}^{L}\Delta t^k\|U^{ck}-u^{ck}\|^2 + C\sum_{k=1}^{L}\Delta t^k\|U^{ik}-u^{ik}\|^2_{L^2(\Gamma)}. \tag{22}$$

After applying the discrete Gronwall's Lemma and (22), we can gain the estimate for $\|Y-Y(u)\|_{l^\infty(J;L^2(\Omega))}$, i.e., the first part of (15).

If we repeat the above manipulation to (18), summing over k from 1 to N_T instead, then we see

$$\|\theta^{N_T}\|^2+2\sum_{k=1}^{N_T}\|\theta^k-\theta^{k-1}\|^2+2\mu\sum_{k=1}^{N_T}\Delta t^k\|\nabla\theta^k\|^2 \le C\sum_{k=1}^{N_T-1}\Delta t^k\|\theta^k\|^2+C\sum_{k=1}^{N_T}\Delta t^k\|U^{ck}-u^{ck}\|^2 + C\sum_{k=1}^{N_T}\Delta t^k\|U^{ik}-u^{ik}\|^2_{L^2(\Gamma)}. \tag{23}$$

Following a similar procedure as the treatment of (22), we can derive the estimate for $\|\nabla(Y-Y(u))\|_{l^2(J;L^2(\Omega))}$ and finish proving (15).

Analogously to the analysis of (15), the result (16) can be established by utilizing (10) and (14), thereby concluding the proof of Lemma 1. □

B. Estimates for U^c-u^c and U^i-u^i

Introduce two integral average operators from $L^2(\Omega)$ (resp. $L^2(\Gamma)$) onto U_h^c (resp. U_h^i), as described in [15] and [27]

$$\Pi_h^s v|_{\tau^s} := \frac{1}{|\tau^s|}\int_{\tau^s} v,\ \forall\ \tau^s\in T_h^s,\ \text{ for } s=c, \text{ or } i$$

and satisfying

$$\begin{cases} \|v-\Pi_h^c v\|_{0,\tau^c}\le Ch_{u^c}\|v\|_{1,\tau^c},\ \forall\ v\in H^1(\Omega), \\ \|v-\Pi_h^i v\|_{0,\tau^i}\le Ch_{u^i}\|v\|_{1,\tau^i},\ \forall\ v\in H^1(\Gamma). \end{cases}$$

Lemma 2 Let $\{y,p,u^c,u^i\}$ be the solution of system (3)-(6), and the corresponding discretized system (9)-(12) have the solution $\{Y,P,U^c,U^i\}$. If $y,p \in L^2(0,T;H^2(\Omega))\cap H^1(0,T;L^2(\Omega)), u^c\in L^2(0,T;H^1(\Omega))$, and $u^i\in L^2(0,T;H^1(\Gamma))$, the following estimate exists:

$$\|U^c-u^c\|_{l^2(J;L^2(\Omega))}+\|U^i-u^i\|_{l^2(J;L^2(\Gamma))} \le C\big\{h_{u^c}+h_{u^i}+\|P(u)-p\|_{l^2(J;H^1(\Omega))}\big\}. \tag{24}$$

Proof. (1). We firstly estimate U^c-u^c. From (5), it follows

$$\begin{aligned}&\alpha_*^c\|U^c-u^c\|^2_{l^2(J;L^2(\Omega))}\\ \le& \sum_{k=1}^{N_T}\sum_{j=1}^{N^c}\Delta t^k\big(\alpha_j^c(u_j^{ck}-U_j^{ck}),u_j^{ck}-U_j^{ck}\big)\\ \le& \sum_{k=1}^{N_T}\sum_{j=1}^{N^c}\Delta t^k\big(\alpha_j^cU_j^{ck},U_j^{ck}-u_j^{ck}\big)\\ &+\sum_{k=1}^{N_T}\sum_{j=1}^{N^c}\Delta t^k\big(B_j^*P^{k-1}(u),U_j^{ck}-u_j^{ck}\big)\\ &+\sum_{k=1}^{N_T}\sum_{j=1}^{N^c}\Delta t^k\Big(B_j^*\big(P^{k-1}(u)-p^{k-1}\big),u_j^{ck}-U_j^{ck}\Big).\end{aligned} \tag{25}$$

Note that $\Pi_h^c u_j^{ck}\in K_h^c$. Then, it follows from (11) that

$$\begin{aligned}&\alpha_*^c\|U^c-u^c\|^2_{l^2(J;L^2(\Omega))}\\ \le& \sum_{k=1}^{N_T}\sum_{j=1}^{N^c}\Delta t^k\Big\{\big(\alpha_j^cU_j^{ck},\Pi_h^cu_j^{ck}-u_j^{ck}\big)\end{aligned}$$

$$+\big(B_j^* p^{k-1}, \Pi_h^c u_j^{ck} - u_j^{ck}\big)$$
$$+\Big(B_j^*(P^{k-1}(u) - p^{k-1}), \Pi_h^c u_j^{ck} - u_j^{ck}\Big)$$
$$+\Big(B_j^*(P^{k-1} - P^{k-1}(u)), \Pi_h^c u_j^{ck} - u_j^{ck}\Big)$$
$$+\Big(B_j^*(P^{k-1}(u) - P^{k-1}), U_j^{ck} - u_j^{ck}\Big)$$
$$+\Big(B_j^*(P^{k-1}(u) - p^{k-1}), u_j^{ck} - U_j^{ck}\Big)\Big\} := \sum_{i=1}^{6} E_i.$$

We now analyze terms from E_1 to E_6 one by one. First, by the definition of Π_h, we observe that

$$E_1 = \sum_{k=1}^{N_T}\sum_{j=1}^{N^c} \Delta t^k \big(\alpha_j^c U_j^{ck}, \Pi_h^c u_j^{ck} - u_j^{ck}\big) = 0, \tag{26}$$

$$\begin{aligned} E_2 =& \sum_{k=1}^{N_T}\sum_{j=1}^{N^c} \Delta t^k \Big(B_j^*\big(p^{k-1} - \Pi_h^c p^{k-1}\big), \Pi_h^c u_j^{ck} - u_j^{ck}\Big) \\ \le& C h_{u^c}^2 \big(\|p\|^2_{l^2(J;H^1(\Omega))} + \|u^c\|^2_{l^2(J;H^1(\Omega))}\big), \end{aligned} \tag{27}$$

$$E_3 \le \varepsilon \|P(u) - p\|^2_{l^2(J;L^2(\Omega))} + C h_{u^c}^2 \|u^c\|^2_{l^2(J;H^1(\Omega))}. \tag{28}$$

Similarly, we obtain

$$E_4 \le \varepsilon \|P - P(u)\|^2_{l^2(J;L^2(\Omega))} + C h_{u^c}^2 \|u^c\|^2_{l^2(J;H^1(\Omega))}, \tag{29}$$

$$E_6 \le C\|P(u) - p\|^2_{l^2(J;L^2(\Omega))} + \varepsilon \|U^c - u^c\|^2_{l^2(J;L^2(\Omega))}. \tag{30}$$

Here, positive constant ε can still be chosen small enough.

(2). We now turn to estimate $U^i - u^i$. The derivation follows a similar way to the estimate of $U^c - u^c$ presented above, and thus we only outline the key steps. It follows from (6) and (12) that

$$\begin{aligned} &\alpha_*^i \|U^i - u^i\|^2_{l^2(J;L^2(\Gamma))} \\ &\le \sum_{k=1}^{N_T}\sum_{j=1}^{N^i} \Delta t^k \Big\{ \big(\alpha_j^i U_j^{ik}, \Pi_h^i u_j^{ik} - u_j^{ik}\big)_{\Gamma_j} \\ &+\big(p^{k-1}, \Pi_h^i u_j^{ik} - u_j^{ik}\big)_{\Gamma_j} \\ &+\Big(P^{k-1}(u) - p^{k-1}, \Pi_h^i u_j^{ik} - u_j^{ik}\Big)_{\Gamma_j} \\ &+\Big(P^{k-1} - P^{k-1}(u), \Pi_h^i u_j^{ik} - u_j^{ik}\Big)_{\Gamma_j} \\ &+\Big(P^{k-1}(u) - P^{k-1}, U_j^{ik} - u_j^{ik}\Big)_{\Gamma_j} \\ &+\Big(P^{k-1}(u) - p^{k-1}, u_j^{ik} - U_j^{ik}\Big)_{\Gamma_j} \\ &+\big(\alpha_j^i \hat{u}_j^{in}, u_j^{ik} - \Pi_h^i u_j^{ik}\big)_{\Gamma_j} \Big\} := \sum_{i=7}^{13} E_i. \end{aligned}$$

Similarly to the analysis of E_1 to E_6, it can be proved that

$$E_7 = \sum_{k=1}^{N_T}\sum_{j=1}^{N^i} \Delta t^k \big(\alpha_j^i U_j^{ik}, \Pi_h^i u_j^{ik} - u_j^{ik}\big)_{\Gamma_j} = 0, \tag{31}$$

$$E_8 \le C h_{u^i}^2 \big(\|p\|^2_{l^2(J;H^1(\Gamma))} + \|u^i\|^2_{l^2(J;H^1(\Gamma))}\big), \tag{32}$$

$$\begin{aligned} E_9 \le& \varepsilon \|P(u) - p\|^2_{l^2(J;H^1(\Omega))} \\ &+ C h_{u^i}^2 \|u^i\|^2_{l^2(J;H^1(\Gamma))}, \end{aligned} \tag{33}$$

$$\begin{aligned} E_{10} \le& \varepsilon \|P - P(u)\|^2_{l^2(J;H^1(\Omega))} \\ &+ C h_{u^i}^2 \|u^i\|^2_{l^2(J;H^1(\Gamma))}, \end{aligned} \tag{34}$$

$$\begin{aligned} E_{12} \le& \varepsilon \|P(u) - p\|^2_{l^2(J;H^1(\Omega))} \\ &+ \varepsilon \|U^i - u^i\|^2_{l^2(J;L^2(\Gamma))}, \end{aligned} \tag{35}$$

$$E_{13} \le C h_{u^i}^2 \big(\|\hat{u}^i\|^2_{l^2(J;H^1(\Omega))} + \|u^i\|^2_{l^2(J;H^1(\Gamma))}\big). \tag{36}$$

(3). For the remaining two terms E_5 and E_{11}, by combining (9)-(10) with (13)-(14), we observe that

$$\begin{aligned} &E_5 + E_{11} \\ &= -\sum_{k=1}^{N_T} \Delta t^k \Big\{ \Big(\frac{\theta^k - \theta^{k-1}}{\Delta t^k}, \zeta^{k-1}\Big) + (\mu \nabla \theta^k, \nabla \zeta^{k-1}) \Big\} \\ &= -\sum_{k=1}^{N_T} \Delta t^k \Big\{ \Big(\frac{\zeta^{k-1} - \zeta^k}{\Delta t^k}, \theta^k\Big) + (\mu \nabla \zeta^{k-1}, \nabla \theta^k) \Big\} \\ &= -\sum_{k=1}^{N_T} \Delta t^k \Big(\alpha^Q \theta^k, \theta^k\Big) \le 0. \end{aligned} \tag{37}$$

Hence, combining the estimates from (25) to (37) and applying (16), we derive

$$\begin{aligned} &\|U^c - u^c\|_{l^2(J;L^2(\Omega))} + \|U^i - u^i\|_{l^2(J;L^2(\Gamma))} \\ &\le C\big\{h_{u^c} + h_{u^i} + \|P(u) - p\|_{l^2(J;H^1(\Omega))}\big\} \\ &+\varepsilon\big\{\|U^c - u^c\|_{l^2(J;L^2(\Omega))} + \|U^i - u^i\|_{l^2(J;L^2(\Gamma))}\big\}. \end{aligned} \tag{38}$$

By choosing ε sufficiently small in (38), we can obtain (24). □

C. *Estimates for $y - Y(u)$ and $p - P(u)$*

For every $\vartheta(t) \in H^2(\Omega)$ and $t \in (0, T]$, we adopt its elliptic projection as $\vartheta_I(t) \in V^h$, which satisfies ([28]-[29]):

$$\big(\mu \nabla(\vartheta - \vartheta_I), \nabla \omega_h\big) = 0, \ \forall \ \omega_h \in V^h. \tag{39}$$

As in [28]- [29], for integer $1 \le r \le 2$, we have

$$\|\vartheta - \vartheta_I\|_{L^2(\Omega)} + h\|\vartheta - \vartheta_I\|_{H^1(\Omega)} \le C h^r \|\vartheta\|_{H^r(\Omega)}. \tag{40}$$

In the subsequent analysis, we will take $\vartheta = y$ and p for consideration, respectively.

Lemma 3 Let systems (3)-(6) and (13)-(14) have the solutions $\{y, p\}$ and $\{Y(u), P(u)\}$, respectively. If the conditions $y,\ p \in L^2(0,T;H^2\ (\Omega)) \cap H^2(0,T;L^2(\Omega)), y_d \in H^1(0,T;L^2(\Omega))$, and $\|y^0 - y_h^0\|_{H^1(\Omega)} \le Ch$ are satisfied. Then, two error estimates hold below:

$$\begin{aligned} &\|y - Y(u)\|_{l^\infty(J;L^2(\Omega))} + \|\nabla(y - Y(u))\|_{l^2(J;L^2(\Omega))} \\ &\qquad \le C\big\{h + \Delta t\big\}, \end{aligned} \tag{41}$$

$$\begin{aligned} &\|p - P(u)\|_{l^\infty(J;L^2(\Omega))} + \|\nabla(p - P(u))\|_{l^2(J;L^2(\Omega))} \\ &\qquad \le C\big\{h + \Delta t\big\}. \end{aligned} \tag{42}$$

Proof. The demonstration of this lemma employs a approach analogous to that used in Lemma 3.3 of [30]. Here, we provide a concise proof of (41). By considering (3) at time $t = t^k$, we observe that:

$$\begin{aligned}&\Big(\frac{y^k - y^{k-1}}{\Delta t^k}, \omega_h\Big) + (\mu \nabla y^k, \nabla \omega_h)\\ &= \sum_{j=1}^{N^c}\Big(B_j u_j^{ck}, \omega_h\Big) + \sum_{j=1}^{N^i}\Big(u_j^{ik}, \omega_h\Big)_{\Gamma_j}\\ &\quad +(\sigma^k, \omega_h),\ \forall\ \omega_h \in V^h,\end{aligned} \tag{43}$$

where $\sigma^k = \dfrac{y^k - y^{k-1}}{\Delta t^k} - \dfrac{\partial y^k}{\partial t}$ is the time truncation error.

By subtracting (13) from (43), selecting $\omega_h = y_I^k - Y^k(u) = \eta^k + y_I^k - y^k$, and applying the projection property (39), we obtain

$$\begin{aligned}&\frac{1}{2\Delta t^k}\Big\{\|\eta^k\|^2 - \|\eta^{k-1}\|^2 + \|\eta^k - \eta^{k-1}\|^2\Big\} + \mu\|\nabla\eta^k\|^2\\ &= \Big(\frac{\eta^k - \eta^{k-1}}{\Delta t^k}, y^k - y_I^k\Big) + (\mu\nabla\eta^k, \nabla(y^k - y_I^k))\\ &\quad + (\sigma^k, y_I^k - Y^k(u)).\end{aligned} \tag{44}$$

Following the same argument as in the derivation of (22), and using (40) and $\|\eta^0\|_1 = \|y^0 - y_h^0\|_1 \le Ch$, it is straightforward to see that

$$\begin{aligned}&\|\eta^L\|^2 + 2\sum_{k=1}^{L}\|\eta^k - \eta^{k-1}\|^2 + 2\mu\sum_{k=1}^{L}\Delta t^k\|\nabla\eta^k\|^2\\ &\le Ch^2 + C(\Delta t)^2 + C\sum_{k=1}^{L-1}\Delta t^k\|\eta^k\|^2.\end{aligned} \tag{45}$$

An immediate application of the discrete Gronwall's lemma to (45) yields the result stated in (41).

Similarly to the proof of (41), the estimate (42) can be established by employing the projection property (39) and the error bound (40). ☐

Combining the estimates provided by Lemmas 1, 2, and 3, we now derive the main result.

Theorem 1 Suppose that system (3)-(6) have the solution $\{y, p, u^c, u^i\}$, and the corresponding discretized system (9)-(12) have the solution $\{Y, P, U^c, U^i\}$. Consequently, under the conditions stated in Lemmas 1, 2, and 3, there hold two bounds as follows:

$$\begin{aligned}&\|u^c - U^c\|_{l^2(J;L^2(\Omega))} + \|u^i - U^i\|_{l^2(J;L^2(\Gamma))}\\ &\le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\},\end{aligned} \tag{46}$$

$$\begin{aligned}&\|y - Y\|_{l^2(J;H^1(\Omega))} + \|p - P\|_{l^2(J;H^1(\Omega))}\\ &\le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\}.\end{aligned} \tag{47}$$

Proof. It follows from (24) and (42) that

$$\begin{aligned}&\|u^c - U^c\|_{l^2(J;L^2(\Omega))} + \|u^i - U^i\|_{l^2(J;L^2(\Gamma))}\\ &\le C\big\{h_{u^c} + h_{u^i} + \|P(u) - p\|_{l^2(J;H^1(\Omega))}\big\}\\ &\le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\}.\end{aligned} \tag{48}$$

This is the estimate (46).

By the triangle inequality and the results of Lemmas 1, 3, and (46), we have

$$\begin{aligned}&\|y - Y\|_{l^\infty(J;L^2(\Omega))} + \|\nabla(y - Y)\|_{l^2(J;L^2(\Omega))}\\ &\le \|y - Y(u)\|_{l^\infty(J;L^2(\Omega))} + \|\nabla(y - Y(u))\|_{l^2(J;L^2(\Omega))}\\ &+ \|Y - Y(u)\|_{l^\infty(J;L^2(\Omega))} + \|\nabla(Y - Y(u))\|_{l^2(J;L^2(\Omega))}\\ &\le C\big\{h + \Delta t + \|U^c - u^c\|_{l^2(J;L^2(\Omega))}\big\}\\ &+ C\|U^i - u^i\|_{l^2(J;L^2(\Gamma))}\\ &\le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\}.\end{aligned} \tag{49}$$

Hence, we have

$$\|y - Y\|_{l^2(J;H^1(\Omega))} \le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\}. \tag{50}$$

Similarly as (50), for $p - P$, we can obtain

$$\|p - P\|_{l^2(J;H^1(\Omega))} \le C\big\{h + h_{u^c} + h_{u^i} + \Delta t\big\}. \tag{51}$$

Combining (50) and (51), we establish the estimate (47), thereby completing the proof. ☐

V. Numerical experiments

In order to demonstrate our theoretical findings derived above, numerical experiments for three distinct scenarios will be carefully performed and comprehensively analyzed. The spatial domain is set as $\Omega = [0,1]\times[0,1]$, and the final time is $T = 1$. Let $\Gamma = \partial\Omega,\ \ \Gamma_1 = \{0\}\times[0,1],\ Q = (0,T)\times\Omega$, $\Sigma = J\times\Gamma$ and $\Sigma_1 = J\times\Gamma_1$.

For simplicity, we discuss a model problem with two controls: one distributed control acting over the domain Ω and another Neumann boundary control applied on the boundary. Specifically, the model problem is:

$$\begin{aligned}\min_{u_1\in K_1,\, u_2\in K_2}&\Big\{J(u_1,u_2) = \frac{1}{2}\int_Q |y(t,x) - y_d(t,x)|^2 dxdt\\ &+\frac{\alpha_1}{2}\int_Q |u_1(t,x)|^2 dxdt + \frac{\alpha_2}{2}\int_{\Sigma_1}|u_2(t,s)|^2 dsdt\Big\}\end{aligned} \tag{52}$$

subject to

$$\begin{cases} y_t(t,x) - \Delta y(t,x)\\ \qquad = f(t,x) + u_1(t,x),\ (t,x)\in Q,\\ \dfrac{\partial y}{\partial n}(t,x) = u_2(t,x),\ (t,x)\in\Sigma_1,\\ y(t,x) = 0,\ (t,x)\in\Sigma\setminus\Sigma_1,\\ y(0,x) = y_0(x),\ x\in\Omega,\end{cases} \tag{53}$$

where K_1 and K_2 are two closed convex sets given below.

According to the system **(OCP-OPT)** mentioned in Section II, the associated co-state equation and optimality conditions of the model (52)-(53) are expressed as follows:

$$\begin{cases} -p_t(t,x) - \Delta p(t,x)\\ \qquad = y(t,x) - y_d(t,x),\ (t,x)\in Q,\\ \dfrac{\partial p}{\partial n}(t,x) = 0,\ (t,x)\in\Sigma_1,\\ p(t,x) = 0,\ (t,x)\in\Sigma\setminus\Sigma_1,\\ p(T,x) = 0,\ x\in\Omega,\end{cases} \tag{54}$$

TABLE I
RESULTS OF ERRORS E_y AND E_p IN CASE I

h	E_y	$Rate$	E_p	$Rate$
1/4	0.840724		1.443046	
1/8	0.441658	0.928702	0.811131	0.831110
1/16	0.223839	0.980467	0.418982	0.953048
1/32	0.112318	0.994868	0.211260	0.987866
1/64	0.056210	0.998687	0.105855	0.996936

and

$$\begin{cases} \int_0^T (\alpha_1 u_1 + p, v - u_1)dt \geq 0, \ \forall\, v \in K_1, \\ \int_0^T (\alpha_2 u_2 + p, w - u_2)_{\Gamma_1} dt \geq 0, \ \forall\, w \in K_2. \end{cases} \tag{55}$$

In our numerical tests, we employ uniform partitions for both $[0, T]$ and Ω. For Ω, we use $I \times I$ uniform meshes, resulting in $2I^2$ triangular elements, where $I = 1/h$ and h denotes the mesh size along the coordinate axes. For simplicity, we set $h = h_{u^c} = h_{u^i} = \Delta t$. Variables y and p are discretized by piecewise linear functions, while variable u_1 and u_2 are approximated by piecewise constant functions. The errors are measured as:

$$\begin{aligned} E_y &= ||y - Y||_{l^2(J;H^1(\Omega))}, \\ E_p &= ||p - P||_{l^2(J;H^1(\Omega))}, \\ E_{u_1} &= ||u_1 - U_1||_{l^2(J;L^2(\Omega))}, \\ E_{u_2} &= ||u_2 - U_2||_{l^2(J;L^2(\Gamma_1))}. \end{aligned} \tag{56}$$

Case I: (**Two Unconstrained Controls**) The parameters are chosen as $\alpha_1 = \alpha_2 = 1$, and the control variables are considered unconstrained, with the admissible sets defined as $K_1 = L^2(0, T; L^2(\Omega))$ and $K_2 = L^2(0, T; L^2(\Gamma_1))$. Based on optimal control theory and (55), the solutions for the two control variables are given by:

$$\begin{cases} u_1(t, x) = -p(t, x), & \forall\, (t, x) \in Q, \\ u_2(t, x) = -p(t, x), & \forall\, (t, x) \in \Sigma_1. \end{cases} \tag{57}$$

To validate the numerical results, two exact solutions are selected as follows:

$$\begin{cases} y(t, x) = (1 - t)\big(\frac{1}{2\pi} sin(2\pi x_1) - x_1^2 + 1\big) sin(2\pi x_2), \\ p(t, x) = (1 - t)\big(cos(2\pi x_1) - x_1^2\big) sin(2\pi x_2). \end{cases} \tag{58}$$

Additionally, we define the corresponding functions $y_d(t, x)$ and $f(t, x)$ such that the governing equations (53)-(54) are satisfied.

Table I presents the results of errors E_y and E_p, while Table II displays the corresponding results of E_{u_1} and E_{u_2}. The results show that the schemes (9)-(12) have one order convergence rate with respect to space and time for all variables, which confirms our theoretical analysis.

Taking the case of $h = \frac{1}{64}$ at time $t = \frac{1}{2}$ for an example, Figures 1 to 4 illustrate the variables y, p, u_1, and u_2, respectively. In each figure, the exact solution is shown on the left, and the approximate solution is illustrated on the right. It is evident that the numerical solutions obtained from the schemes (9)-(12) closely approximate the exact solutions, demonstrating the effectiveness and accuracy of the proposed methods.

TABLE II
RESULTS OF ERRORS E_{u_1} AND E_{u_2} IN CASE I

h	E_{u_1}	$Rate$	E_{u_2}	$Rate$
1/4	0.328547		0.387827	
1/8	0.149901	1.132085	0.174530	1.151942
1/16	0.069135	1.116525	0.082288	1.084715
1/32	0.033586	1.041544	0.040361	1.027724
1/64	0.016660	1.011449	0.020076	1.007522

TABLE III
RESULTS OF ERRORS E_y AND E_p IN CASE II

h	E_y	$Rate$	E_p	$Rate$
1/4	0.840530		1.443048	
1/8	0.441580	0.928625	0.811131	0.831112
1/16	0.223827	0.980289	0.418982	0.953048
1/32	0.112317	0.994811	0.211260	0.987866
1/64	0.056210	0.998672	0.105855	0.996936

TABLE IV
RESULTS OF ERRORS E_{u_1} AND E_{u_2} IN CASE II

h	E_{u_1}	$Rate$	E_{u_2}	$Rate$
1/4	0.227762		0.387831	
1/8	0.105213	1.114215	0.174531	1.151940
1/16	0.048921	1.104783	0.082289	1.084723
1/32	0.023818	1.038417	0.040361	1.027728
1/64	0.011806	1.012462	0.020076	1.007523

Case II: (**One Constrained Control**) We impose constraints on the control variable u_1, defining the admissible sets as $K_1 = \{v \in L^2(0, T; L^2(\Omega)),\ v \geq 0\}$ and $K_2 = L^2(0, T; L^2(\Gamma_1))$. Additionally, we set $\alpha_1 = \alpha_2 = 1$. Based on optimal control theory and (55), the solutions for the two control variables are given by:

$$\begin{cases} u_1(t, x) = \max\{0, -p(t, x)\}, & \forall\, (t, x) \in Q, \\ u_2(t, x) = -p(t, x), & \forall\, (t, x) \in \Sigma_1. \end{cases} \tag{59}$$

For this case, we retain the same exact state and co-state solutions as specified in (58). Under the conditions of (58) and (59), we can determine the corresponding functions $y_d(t, x)$ and $f(t, x)$ that satisfy the equations (53)-(54).

It should be emphasized that the fully discrete scheme (10) for the co-state variable is time-backward and coupled with the state variable scheme (9). Due to the constrained control variable u_1, the schemes (9)-(12) cannot be solved directly. To overcome this, we employ an iterative method, which not only simplifies matrix computations but also reduces computational complexity in time. This approach has been successfully applied in previous studies, such as [31]- [34]. Tables III and IV present the results of errors E_y, E_p, E_{u_1}, and E_{u_2}, respectively. The results demonstrate that the schemes (9)-(12) achieve a first-order convergence rate with respect to space and time for all variables, further confirming the validity of our theoretical analysis.

Taking the case of $h = \frac{1}{64}$ at time $t = \frac{1}{2}$ for an example, Figures 5 to 8 illustrate the variables y, p, u_1, and u_2, respectively. The results demonstrate that the numerical solutions obtained from the schemes (9)-(12) approximate the exact solutions well.

Case III: (**Effect of Regularization Parameters**) We consider the influence of varying regularization parameters

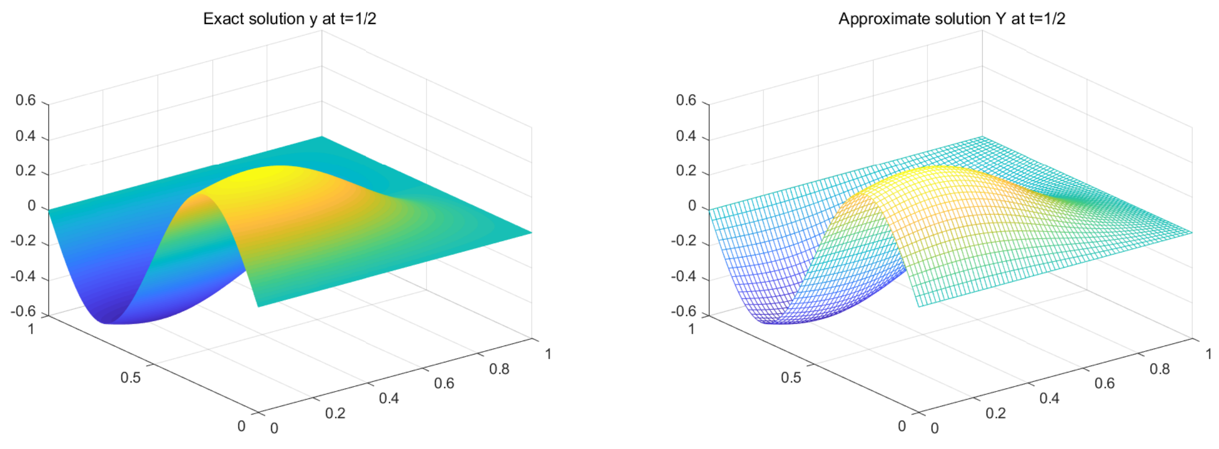

Fig. 1. The state variable y in Case I

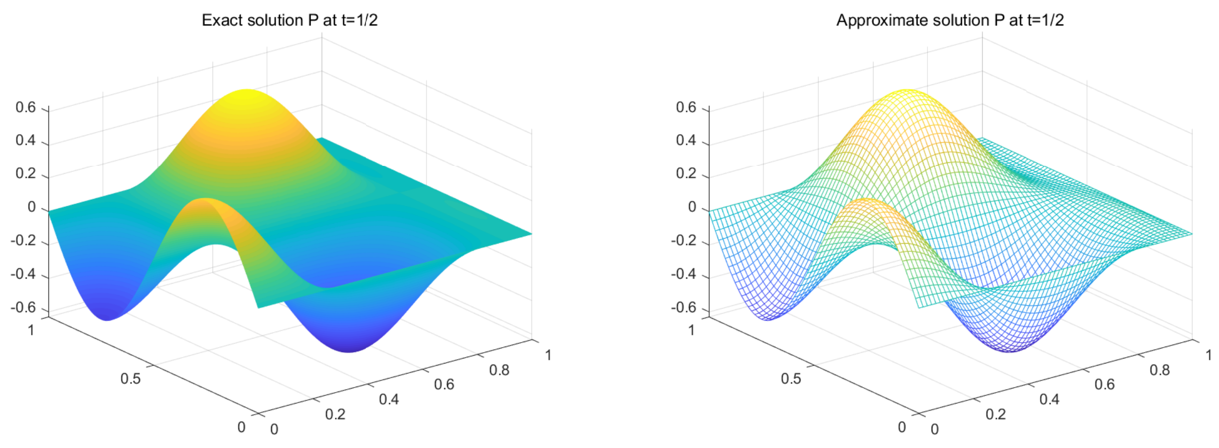

Fig. 2. The co-state variable p in Case I

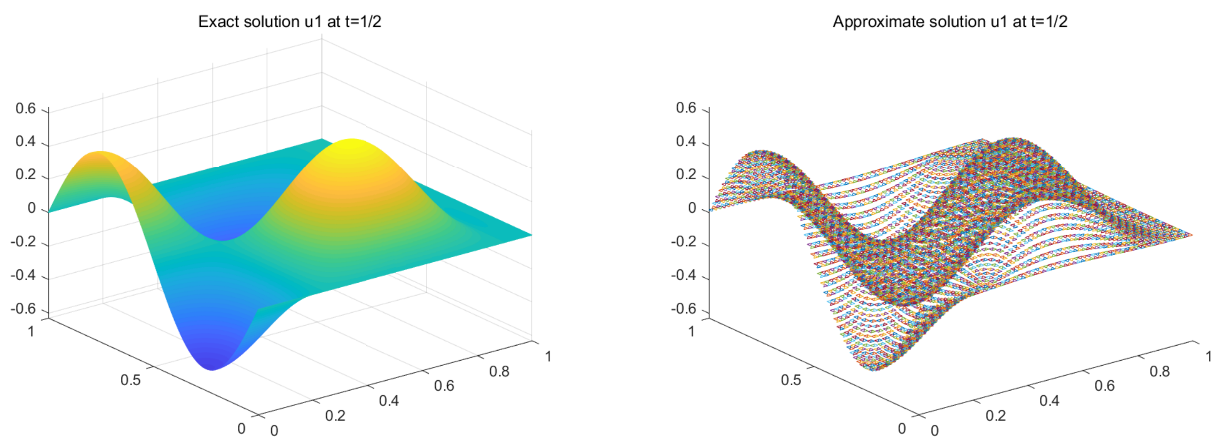

Fig. 3. The control variable u_1 in Case I

α_1 and α_2 used in Case II. We select the following exact state and co-state solutions

$$\begin{cases} y(t,x) = \dfrac{1}{\alpha_2}(1-t)\big(\dfrac{1}{2\pi}sin(2\pi x_1) - x_1^2 + 1\big)sin(2\pi x_2), \\ p(t,x) = (1-t)\big(cos(2\pi x_1) - x_1^2\big)sin(2\pi x_2). \end{cases} \tag{60}$$

Then, by the optimal control theory and (55), we know that the two control solutions are

$$\begin{cases} u_1(t,x) = \max\{0, -\frac{1}{\alpha_1}p(t,x)\}, \ \forall\ (t,x) \in Q, \\ u_2(t,x) = -\frac{1}{\alpha_2}p(t,x), \ \forall\ (t,x) \in \Sigma_1. \end{cases} \tag{61}$$

And, under the conditions specified in (61) and (60), we can determine the corresponding functions $y_d(t,x)$ and $f(t,x)$ that satisfy the equations (53)-(54).

Table V shows how the number of iterations varies with

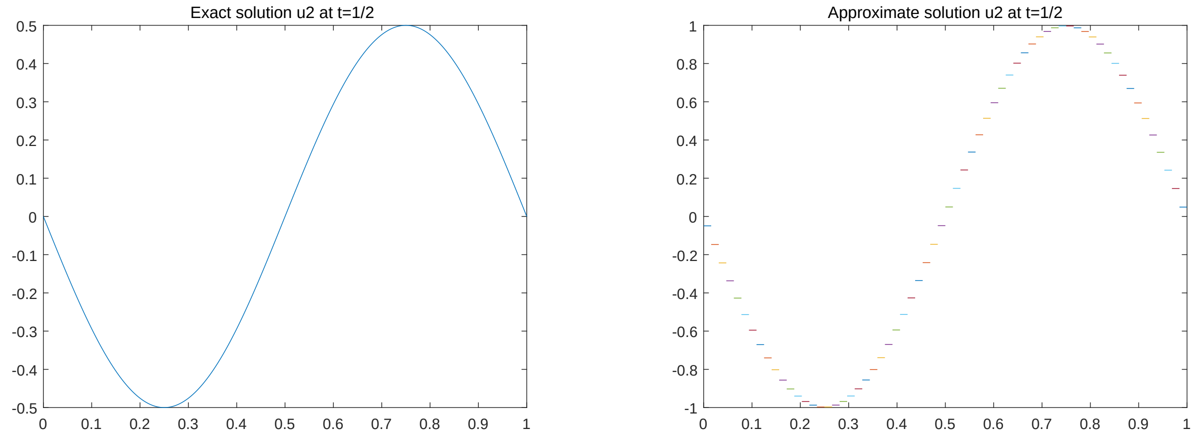

Fig. 4. The control variable u_2 in Case I

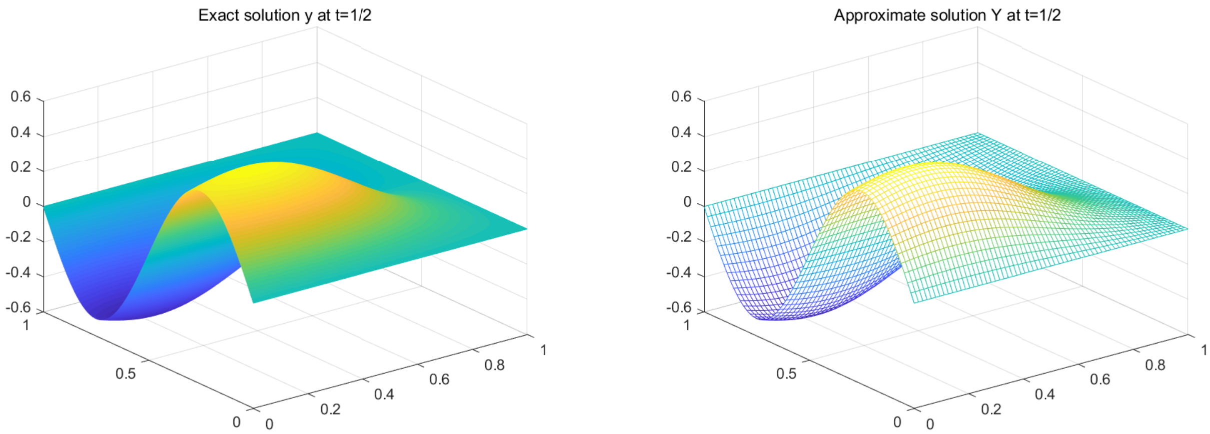

Fig. 5. The state variable y in Case II

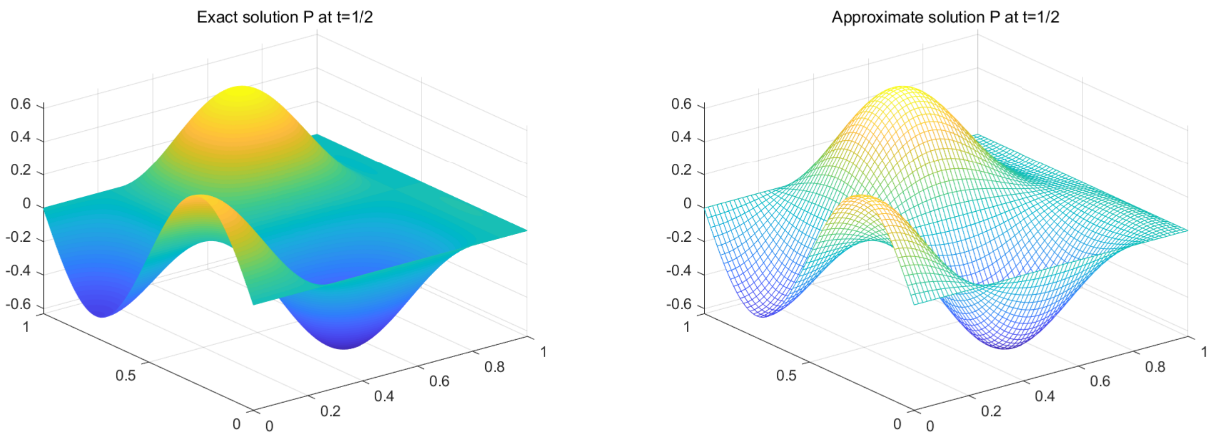

Fig. 6. The co-state variable p in Case II

different α_1 and α_2 under the same stopping tolerance. It is observed that the iteration number increases as either α_1 or α_2 decreases. Furthermore, when α_1 or α_2 decreases below a certain threshold, the iterative method fails to converge. This indicates that continuously reducing the regularization parameters α_1 and α_2 does not ensure convergence, highlighting the need for a balance between these parameters. Figure 9 illustrates the decay of the value of the discretized functional $J_h(U_1, U_2)$ for the case $\alpha_1 = \alpha_2 = 0.1$. Here, $J_h(U_1, U_2)$ is defined as in (7). The plot reveals a sharp decline in the functional value at the initial stages of the iteration, followed by a much slower decrease as the iteration progresses. This behavior demonstrates the effectiveness of our algorithm in optimizing the objective functional.

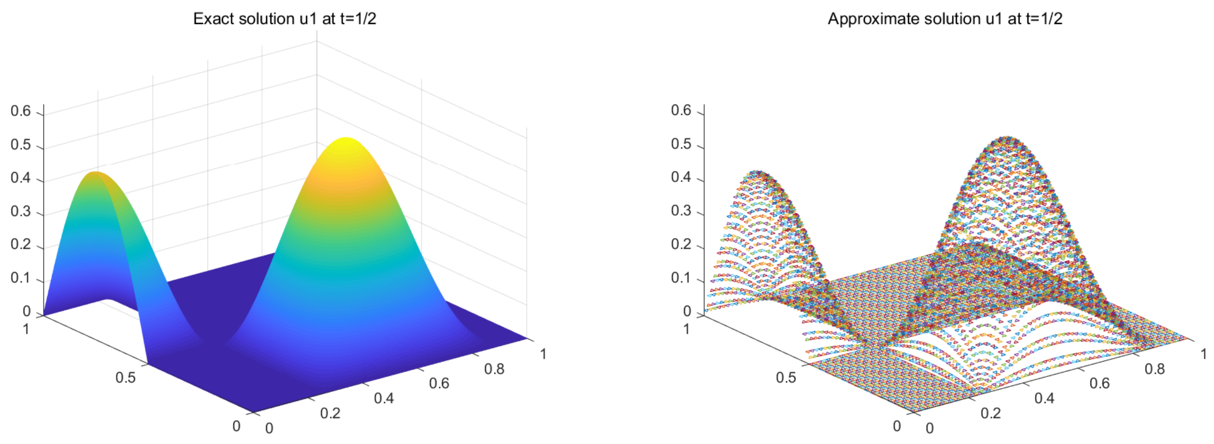

Fig. 7. The control variable u_1 in Case II

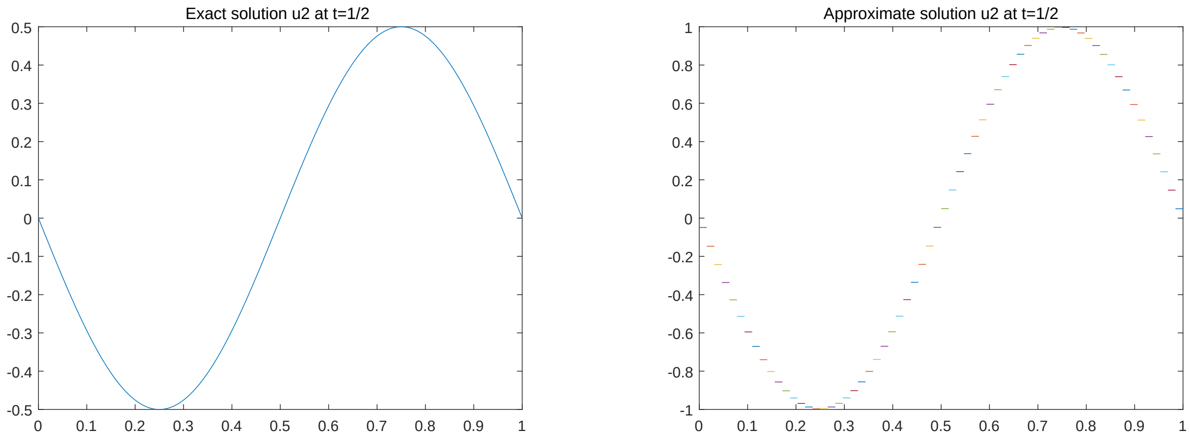

Fig. 8. The control variable u_2 in Case II

TABLE V
ITERATION NUMBER WITH DIFFERENT α_1 AND α_2

α_1	α_2	*Iteration number*
1	1	4
0.1	1	6
0.01	1	11
0.001	1	∞
1	0.1	7
1	0.01	∞
0.1	0.1	8
0.01	0.01	∞

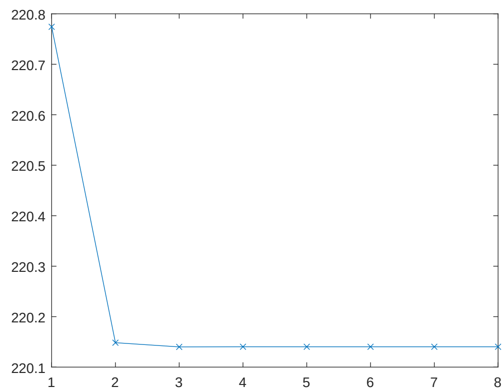

Fig. 9. The value of $J_h(U_1, U_2)$ decays with iteration

VI. CONCLUSIONS

We have investigated the whole process of utilizing FEM to solve an optimization problem constrained by heat conduction equation, involving two types of controls: one acting in the domain and the other on its boundary. Through applying the theory for PDEs-constrained optimal control problem, we obtain the corresponding co-state equation and optimality conditions, which, together with the state equation, constitute the optimality system. For the purpose of establishing the fully discrete approximate schemes, the state and co-state variables are approximated using piecewise linear continuous functions, whereas piecewise constant functions are employed for the control variables. Through rigorous theoretical analysis, we obtain a priori error estimates in appropriate norms for all considered variables. Furthermore, we conduct numerical experiments for three distinct scenarios to validate and illustrate our theoretical findings.

REFERENCES

[1] J.L. Lions, Optimal control of systems governed by partial differential equations, Springer-Verlag, Berlin, 1971.

[2] W.B. Liu and N.N. Yan, Adaptive finite element method for optimal control governed by PDEs, Science Press, Beijing, 2008.
[3] F. Tröltzsch, Optimal control of partial differential equations: theory, methods, and applications, American Mathematical Society, Providence, Rhode Island, 2010.
[4] D. Meidner and 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., vol. 47, no. 3, pp. 1150-1177, 2008.
[5] D. Meidner and 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., vol. 47, no. 3, pp. 1301-1329, 2008.
[6] H.F. Fu and H.X. Rui, "A prior error for optimal control problem governed by transient advection-diffusion equations," J. Sci. Comput., vol. 38, no. 3, pp. 290-315, 2009.
[7] H.F. Fu and H.X. Rui, "A characteristic-mixed finite element method for time-dependent convection-diffusion optimal control problem," Appl. Math. Comput., vol. 218, no. 7, pp. 3430-3440, 2011.
[8] Y.P. Chen and Z.L. Lu, "Error estimates of fully discrete mixed finite element methods for semilinear quadratic parabolic optimal control problem," Comput. Methods Appl. Mech. Engrg., vol. 199, no. 23-24, pp. 1415-1423, 2010.
[9] H.F. Fu, H. Guo, J. Hou and J.L. Zhao, "A priori error analysis of stabilized mixed finite element method for reaction-diffusion optimal control problems," Bound. Value Probl., article number: 23, 2016.
[10] Y.L. Tang and Y.C. Hua, "Superconvergence of semidiscrete finite element methods for bilinear parabolic optimal control problems," J. Inequal. Appl., article number: 62, 2017.
[11] T.J. Sun, "Discontinuous Galerkin finite element method with interior penalties for convection diffusion optimal control problem," Int. J. Numer. Anal. Model., vol. 7, no. 1, pp. 87-107, 2010.
[12] T.J. Sun, L. Ge and W.B. Liu, "Equivalent a posteriori error estimates for a constrained optimal control problem governed by parabolic equations," Int. J. Numer. Anal. Model., vol. 10, pp. 1-23, 2013.
[13] Z.L. Lu, S.H. Du and Y.L. Tang, "New a posteriori error estimates of mixed finite element methods for quadratic optimal control problems governed by semilinear parabolic equations with integral constraint," Bound. Value Probl., article number: 230, 2013.
[14] W.B. Liu and N.N. Yan, "A posteriori error estimates for distributed convex optimal control problems," Advance. Comput. Math., vol. 15, no. 1-4, pp. 285-309, 2001.
[15] W.B. Liu and N.N. Yan, "A posteriori error estimates for control problems governed by Stokes equations," SIAM J. Numer. Anal., vol. 40, no. 5, pp. 1850-1869, 2002.
[16] R. Becker and B. Vexler, "A posteriori error estimation for finite element discretization of parameter identification problems," Numer. Math., vol. 96, no. 3, pp. 435-459, 2004.
[17] W. Gong and N.N. Yan, "Adaptive finite element method for elliptic optimal control problems: convergence and optimality," Numer. Math., vol. 135, no. 4, pp. 1121-1170, 2017.
[18] A. Foucquier, S. Robert, F. Suard, L. Stephan and A. Jay, "State of the art in building modelling and energy performance prediction: a review," Renew. Sustain. Energy Rev., vol. 23, pp. 272-288, 2013.
[19] S. Sunder and J. Dion, Federal research and development agenda for net-zero energy, high-performance green buildings, Tech. rep., National Science and Technology Council, Committee on Technology, Subcommittee on Buildings Technology Research and Development, 2008.
[20] F. Bachmann, D. Beermann, J.J. Lu and S. Volkwein, "POD-based mixed-integer optimal control of the heat equation," J. Sci. Comput., vol. 81, no. 1, pp. 48-75, 2019.
[21] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke and A. Mahajan, "Mixed-integer nonlinear optimization," Acta Numer., vol. 22, pp. 1-131, 2013.
[22] S. Leyffer, P. Cay, D. Kouri and B. van Bloemen Waanders, Mixed-integer PDE-constrained optimization, Tech. rep., Oberwolfach Report, 2015.
[23] K. Kunisch and S. Volkwein, "Galerkin proper orthogonal decomposition methods for parabolic problems," Numer. Math., vol. 90, pp. 117-148, 2001.
[24] M. Hinze and S. Volkwein, "Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition," Comput. Optim. Appl., vol. 39, no. 3, pp. 319-345, 2008.
[25] F. Tröltzsch and S. Volkwein, "POD a-posteriori error estimates for linear-quadratic optimal control problems," Comput. Optim. Appl., vol. 44, no.1, pp. 83-115, 2009.
[26] Z.D. Luo, J. Chen and I.M. Navon, "Mixed finite element formulation and error estimates based on proper orthogonal decomposition for the non-stationary Navier-Stokes equations," SIAM J. Numer. Anal., vol. 47, no.1, pp. 1-19, 2008.
[27] P.G. Ciarlet, The finite element method for elliptic problems, SIAM, Philadelphia, 2002.
[28] V. Thomée, Galerkin finite element methods for parabolic problems, Springer, Springer Series in Computational Mathematics, Berlin, 1997.
[29] M.F. Wheeler, "A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations," SIAM J. Numer. Anal., vol. 10, no. 4, pp. 723-759, 1973.
[30] R.R. Zheng and T.J. Sun, "A priori error estimates of finite element methods for an optimal control problem governed by a one-prey and one-predator model," Journal of Shandong University (Natural Science), vol. 55, no. 1, pp. 23-31, 2020.
[31] W.Y. Liu and K.Y. Ma, "An iterative non-overlapping domain decomposition method for optimal boundary control problems governed by parabolic equations," IAENG International Journal of Applied Mathematics, vol. 46, no. 3, pp. 291-297, 2016.
[32] F.F. Du and T.J. Sun, "A cubic B-spline finite element method for optimal control problem governed by nonlinear parabolic equation describing crystal surface growth," IAENG International Journal of Applied Mathematics, vol. 52, no. 3, pp. 754-759, 2022.
[33] C.J. Yang and T.J. Sun, "Crank-Nicolson finite difference schemes for parabolic optimal Dirichlet boundary control problems," Math. Methods Appl. Sci., vol. 45, no. 12, pp. 7346-7363, 2022.
[34] C.J. Yang and T.J. Sun, "BDF2 schemes for optimal parameter control problems governed by bilinear parabolic equations," Optim. Contr. Appl. Met., vol. 44, no. 4, pp. 2055-2081, 2023.