Hamburger Beiträge
zur Angewandten Mathematik

A priori Error Analysis for Finite Element Approximation of Parabolic Optimal Control Problems with Pointwise Control

Wei Gong, Michael Hinze, and Zhaojie Zhou

Nr. 2011-07
April 2011

A PRIORI ERROR ANALYSIS FOR FINITE ELEMENT APPROXIMATION OF PARABOLIC OPTIMAL CONTROL PROBLEMS WITH POINTWISE CONTROL

WEI GONG *, MICHAEL HINZE †, AND ZHAOJIE ZHOU◇

Abstract: We consider finite element approximations of parabolic control problems with pointwise control in this paper. The state equation exhibits low regularity due to the control imposed pointwisely. To discretize the optimal control problem we use variational discretization together with piecewise linear and continuous finite elements for the space discretization of the state, and the dG(0) method for time discretization. We prove a priori error estimates for control, state and adjoint state. Numerical experiments are provided which confirm the theoretical results.

Key words. optimal control problem, finite element method, a priori error estimate, parabolic equation, pointwise control.

Subject Classification: 49J20, 49K20, 65N15, 65N30.

1. Introduction

Optimal control problems with time-dependent control play an important role in many applications, so that the numerical treatment of these problems currently is a hot research topic. For an overview concerning a priori error analysis for parabolic optimal control problems we refer to e.g. [25], a posteriori error analysis can be found in e.g. [20]. A priori error analysis for problems with state constraints can be found in, e.g., [7], [14] and [24].

However, all the work mentioned above are focused on distributed control, namely, the control acts on parts of the domain, so that for both control constrained and state constrained problems the state equations all exhibit standard regularities. In many applications, however, control can only act locally at finitely points of the domain, which is called pointwise control. In this case, the state equation has limited regularity, so it is of primal interest to investigate the behavior of these kind of control problems both theoretically and numerically.

There are a lot of applications for control problems where governing state equation exhibit low regularity resulted from the pointwise control. For example, by controlling the magnitude of some pointwise heating source we can achieve a desired temperature distribution with as less as possible energy consumption. There are also applications in inverse problems. For the data assimilation problems in environment surveillance, one needs to determine the magnitudes of pollutant emission from some given locations with observations, the regularized problem can be treated and analyzed in the content of pointwise control problems. Another typical application in this context concerns control of pollution sources in industrial areas, ensuring required air and/or water quality in the living areas of the population, see Section 5 for details.

For optimal control problems with pointwise control, Dean, Gubernatis and Ramos etal. studied such problems with governing equations of Burger's type in [6] and [26], respectively. Chryssoverghi studied approximation methods for optimal pointwise control of parabolic systems in [3]. Droniou

Date: April 14, 2011.

* LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: wgong@lsec.cc.ac.cn.

† Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146, Hamburg, Germany Email: michael.hinze@uni-hamburg.de.

◇ School of Mathematics Sciences, Shandong Normal University, 250014 Ji'nan, China. Email: zzj534@amss.ac.cn .

and Raymond also analyzed the optimal pointwise control of semilinear parabolic equations in [9]. However, there seems to be no contributions to the finite element analysis of such kind of problems.

The aim of this paper is to analyze the finite element approximation of parabolic optimal control problems with pointwise control. These problems are difficult due to the low regularity of state equation through the pointwise control. We use standard piecewise linear and continuous finite elements for the space discretization of the state, while the dG(0) method is used for time discretization. Based on the error estimates for finite element approximation of parabolic equations with measure data in space presented in [13], a priori error estimates for control, state and adjoint state are derived. Numerical experiments are also provided to confirm our theoretical results.

The outline of the paper is as follows. In Section 2 we present the parabolic optimal control problem with pointwise control. Existence and uniqueness as well as regularity results for the solutions are established. In Section 3 we present the fully discrete finite element approximation for the state equation and corresponding optimal control problems together with the stability estimate for the discrete scheme. Section 4 is devoted to the a priori estimates for the discretization error of state equation and thus a priori error analysis for optimal control problems. We also give some extensions to problems with states governed by the time-dependent convection diffusion equations and additional state constraints. We finally carry out some numerical experiments to confirm our theoretical findings.

For a convex polygonal domain $\Omega\subset\mathbb{R}^d$, $d=2$ or 3, we adopt the standard notation $W^{m,s}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{m,s,\Omega}$ and seminorm $|\cdot|_{m,s,\Omega}$. We denote by $H^m(\Omega)$ with norm $\|\cdot\|_{m,\Omega}$ and seminorm $|\cdot|_{m,\Omega}$ for $s=2$. Note that $H^0(\Omega)=L^2(\Omega)$ and $H_0^1(\Omega)=\{v\in H^1(\Omega);\ v=0 \text{ on } \partial\Omega\}$. We denote by $L^r(0,T;W^{m,s}(\Omega))$ the Banach space of all L^r integrable functions from $[0,T]$ into $W^{m,s}(\Omega)$ with norm $\|v\|_{L^r(0,T;W^{m,s}(\Omega))}=\Big(\int_0^T\|v\|^r_{m,s,\Omega}dt\Big)^{\frac{1}{r}}$ for $1\leqslant r<\infty$, and the standard modification for $r=\infty$. We set $W(0,T):=L^2(0,T;H_0^1(\Omega))\cap H^1(0,T;H^{-1}(\Omega))$ and $X:=L^2(0,T;H^2(\Omega)\cap H_0^1(\Omega))\cap H^1(0,T;L^2(\Omega))\hookrightarrow C([0,T];H^1(\Omega))$. In addition, c and C denote generic positive constants.

2. Optimal control problem

Let Ω be an open bounded domain in $\mathbb{R}^d$ ($d=2$ or 3) with boundary $\Gamma=\partial\Omega$, $T>0$, $\Omega_T=\Omega\times(0,T]$, $\Gamma_T=\partial\Omega\times(0,T]$. We consider the following parabolic problem

$$
\begin{cases}
y_t+Ay=\displaystyle\sum_{i=1}^m u_i(t)\delta_{X_i} & \text{in } \Omega_T,\\
y(x,t)=0 & \text{on } \Gamma_T,\\
y(x,0)=y_0(x) & \text{in } \Omega,
\end{cases}
\tag{2.1}
$$

where $y:\Omega_T\to\mathbb{R}$ is the state variable. The control $u_i(t)$ only depends on time and acts on finitely many points $X_i\in\Omega$, δ_{X_i} represents the Dirac measure concentrated at X_i, $i=1,2,\cdots,m$ and $y_0\in L^2(\Omega)$ is a given function.

We denote $\mathcal{M}(\bar\Omega)$ the space of the real and regular Borel measures on Ω, which can be defined as the dual space of $\mathcal{C}(\bar\Omega)$ with its natural norm

$$\|\mu\|_{\mathcal{M}(\bar\Omega)}=\sup\Big\{\int_\Omega vd\mu:\ v\in\mathcal{C}(\bar\Omega)\text{ and }\|v\|_{\mathcal{C}(\bar\Omega)}\leqslant 1\Big\}.$$

We denote the L^2-inner products on $L^2(\Omega)$ and $L^2(\Omega_T)$ by

$$(v,w)=\int_\Omega vwdx\quad\forall\ v,w\in L^2(\Omega)$$

and

$$(v,w)_{\Omega_T}=\int_{\Omega_T}vwdxdt\quad\forall\ v,w\in L^2(\Omega_T),$$

respectively.

The operator A is assumed to be a second order elliptic partial differential operator of the form

$$Ay = -\sum_{i,j=1}^{d} \partial_{x_j}(a_{ij}\partial_{x_i} y) + a_0 y,$$

where $a_0 \in L^\infty(\Omega)$, $a_0(x,t) \geqslant 0$ for all $(x,t) \in \Omega_T$, a_{ij} $(1 \leqslant i,j \leqslant d)$ is Lipschitz continuous on Ω_T and satisfies the following uniform ellipticity condition:

$$\sum_{i,j=1}^{d} a_{ij}\xi_i\xi_j \geqslant c|\xi|^2, \quad c > 0, \quad \forall\, \xi \in \mathbb{R}^d, \quad x \in \Omega.$$

Furthermore, $\partial_{\boldsymbol{n}_A} = \sum\limits_{i,j=1}^{n} a_{ij}\partial_{x_j} n_i$ with $\boldsymbol{n}$ denoting the unit outer normal to Γ. We will denote by A^* the adjoint operator of A:

$$A^* y = -\sum_{i,j=1}^{d} \partial_{x_j}(a_{ji}\partial_{x_i} y) + a_0 y.$$

Thus we can define the following bilinear forms associated with A on Ω and Ω_T:

$$a(v,w) = \int_\Omega (\sum_{i,j=1}^{d} a_{ij}\partial_{x_i} v \partial_{x_j} w + a_0 v w) dx \quad \forall\, v, w \in H_0^1(\Omega)$$

and

$$a(v,w)_{\Omega_T} = \int_{\Omega_T} (\sum_{i,j=1}^{d} a_{ij}\partial_{x_i} v \partial_{x_j} w + a_0 v w) dxdt \quad \forall\, v, w \in L^2(0,T; H_0^1(\Omega)).$$

We set $U := L^2(0,T;\mathbb{R}^m)$ and define the linear, bounded control operator $B : U \to L^2(0,T;\mathcal{M}(\bar{\Omega}))$ by

$$(Bu)(x,t) := \sum_{i=1}^{m} u_i(t)\delta_{X_i} \quad t \in [0,T], \;\; X_i \in \Omega. \tag{2.2}$$

The weak solution of problems (2.1) can be defined by transposition techniques (see Lions and Magenes [19]). We have following results concerning existence and uniqueness, as well as on regularity of solutions to problem (2.1).

Theorem 2.1. *Problem (2.1) admits a unique solution $y \in L^2(0,T;L^2(\Omega))$ in the sense that*

$$-(y, \partial_t v)_{\Omega_T} + (y, A^* v)_{\Omega_T} = \langle Bu, v\rangle_{\Omega_T} + (y_0, v(\cdot,0)) \tag{2.3}$$

for any $v \in W(\Omega)$, where

$$W(\Omega) = \big\{v \in X : \;\; \partial_t v + A^* v \in L^2(\Omega_T) \;\; \text{and } \partial_{\boldsymbol{n}_{A^*}} v \in L^2(\Gamma_T), \; v(\cdot,T) = 0\big\}.$$

Here

$$\langle Bu, v\rangle_{\Omega_T} = \int_{\bar{\Omega}_T} Buv dxdt = \int_0^T \sum_{i=1}^{m} u_i(t) v(X_i,t) dt \quad \forall\, v \in L^2(0,T;\mathcal{C}(\bar{\Omega})).$$

Furthermore, there exists a positive constant C only depending on data, such that

$$\|y\|_{L^2(0,T;L^2(\Omega))} \leqslant C\big(\|u\|_{L^2(0,T;\mathbb{R}^m)}(\sum_{i=1}^{m} \|\delta_{X_i}\|_{\mathcal{M}(\bar{\Omega})}) + \|y_0\|_{0,\Omega}\big). \tag{2.4}$$

Moreover, we have $y \in L^1(0,T;W^{1,s}(\Omega)) \cap \mathcal{C}(0,T;W^{1,s'}(\Omega)')$ and $\partial_t y \in L^1(0,T;W^{1,s'}(\Omega)')$ with

$$\|y\|_{L^1(0,T;W^{1,s}(\Omega))} \leqslant C\big(\|u\|_{L^2(0,T;\mathbb{R}^m)}(\sum_{i=1}^{m} \|\delta_{X_i}\|_{\mathcal{M}(\bar{\Omega})}) + \|y_0\|_{0,\Omega}\big), \tag{2.5}$$

where $s \in [1, \frac{d}{d-1})$ and s' is the conjugate number of s such that $\frac{1}{s} + \frac{1}{s'} = 1$.

Proof. For the proof we refer to, e.g., [2], [13] and [23]. □

Remark 2.2. *If $d = 1$, the function $u(t) \cdot \delta(x)$ belongs to $L^2(0,T; H^{-1}(\Omega))$, this property implies in turn that(see [12])*

$$y \in L^2(0,T; H^1(\Omega)) \cap C(0,T; L^2(\Omega)), \quad \frac{\partial y}{\partial t} \in L^2(0,T; H^{-1}(\Omega)).$$

However, we will not exploit these special situations and consider the cases $d = 2$ or 3.

We consider the parabolic optimal control problem

$$\left\{ \begin{aligned} \min_{u \in U_{ad}} \ J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega_T)} + \frac{\alpha}{2}\|u\|^2_{L^2(0,T;\mathbb{R}^m)} \\ y \text{ be the solution of problem (2.1),} \end{aligned} \right. \tag{2.6}$$

where $y_d \in L^2(0,T; L^2(\Omega))$ is a given function and $\alpha > 0$. The admissible control set is given by

$$U_{ad} := \left\{ u \in L^2(0,T;\mathbb{R}^m) : \quad a_i \leqslant u_i(t) \leqslant b_i, \quad i = 1, 2, \cdots, m \quad a.e. \ in \quad (0,T) \right\}, \tag{2.7}$$

where $a_i < b_i$ are given constants.

Since the set of admissible controls is closed and convex one obtains the existence of a unique solution $u \in U_{ad}$ to problem (2.6) by standard arguments (see [18]). Moreover, we have the following first order optimality condition:

Theorem 2.3. *Assume that $u \in U_{ad}$ is the solution of problem (2.6) and let y be the corresponding state given by (2.3). Then there exists a unique adjoint state $p \in W(0,T)$ satisfying*

$$\left\{ \begin{aligned} -p_t + A^* p &= y - y_d \quad in \ \Omega_T, \\ p(x,t) &= 0 \quad on \ \Gamma_T, \\ p(x,T) &= 0 \quad in \ \Omega. \end{aligned} \right. \tag{2.8}$$

Furthermore,

$$\sum_{i=1}^m \int_0^T (\alpha u_i(t) + p(X_i,t))(\tilde{v}_i - u_i) dt \geqslant 0 \quad \forall \ \tilde{v} \in U_{ad}. \tag{2.9}$$

Let

$$P_{U_{ad}}(u_i(t)) = \max(a_i, \min(b_i, u_i(t))), \quad i = 1, 2, \cdots, m$$

denote the pointwise projection onto the admissible set U_{ad}. Then the optimality condition (2.9) is equivalent to

$$u_i(t) = P_{U_{ad}}(-\frac{1}{\alpha} p(X_i,t)), \quad i = 1, 2, \cdots, m. \tag{2.10}$$

From (2.10) we deduce the regularity results summarized in the next theorem.

Theorem 2.4. *Assume that $u \in U_{ad}$ is the solution of problem (2.6), y is the associated state and p is the adjoint state, then we have*

$$y \in L^1(0,T; W_0^{1,s}(\Omega)) \cap L^2(0,T; L^2(\Omega)) \cap \mathcal{C}(0,T; W^{1,s'}(\Omega)'),$$
$$p \in L^2(0,T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0,T; L^2(\Omega)), \quad u \in L^\infty(0,T;\mathbb{R}^m).$$

Proof. From Theorem 2.1 we have $y \in L^1(0,T; W_0^{1,s}(\Omega)) \cap L^2(0,T; L^2(\Omega)) \cap \mathcal{C}(0,T; W^{1,s'}(\Omega)')$, then standard regularity results for parabolic equation give $p \in L^2(0,T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0,T; L^2(\Omega))$. Thus, $p \in L^2(0,T; \mathcal{C}(\bar{\Omega}))$ by an embedding theorem, so that control constraints (2.7) and property (2.10) impliy the stated regularity for optimal control u. □

Remark 2.5. *We can only conclude that the control u belongs to* $L^\infty(0,T,\mathbb{R}^m)$ *for constrained case, which is dominated by the pointwise regularity in space of the adjoint state p. It seems to be unable to improve the results due to the loss of regularity of state y, and thus the adjoint state p, even for the unconstrained case.*

3. Finite element approximation of optimal control problem

The goal of this section is to investigate the fully discrete finite element approximation of the state equation and thus the optimal control problems (2.6), and to derive the corresponding stability estimates.

To define the fully discrete finite element discretization scheme to (2.6) we consider a family of regular triangulation $\mathscr{T}^h$ of $\bar\Omega$, such that $\bar\Omega = \cup_{\tau\in\mathscr{T}^h}\bar\tau$. Let $h = \max_\tau h_\tau$, where h_τ denotes the diameter of the element τ. Moreover, we suppose that $\mathscr{T}^h$ is quasi-uniform. Associated with $\mathscr{T}^h$ is a finite dimensional subspace V^h of $C(\bar\Omega)$, which consists of piecewise linear polynomials. We set $V_0^h = V^h \cap H_0^1(\Omega)$. Since $\mathscr{T}^h$ is quasi-uniform, the following inverse estimates (see [4])

$$\|v_h\|_{s,\Omega} \leqslant Ch^{l-s}\|v_h\|_{l,\Omega},\quad 0\leqslant l\leqslant s\leqslant 1, \tag{3.1}$$

$$\|v_h\|_{0,\infty,\Omega} \leqslant Ch^{-\frac{d}{2}}\|v_h\|_{0,\Omega} \tag{3.2}$$

hold for all $v_h \in V^h$.

Let $Q_h : L^2(\Omega) \to V_h$ be the L^2 projection operator defined by

$$(Q_h v, w_h) = (v, w_h)\quad \forall\, w_h \in V^h \tag{3.3}$$

and $R_h : H_0^1(\Omega) \to V_0^h$ be the Ritz projection operator given by

$$a(R_h v, w_h) = a(v, w_h)\quad \forall\, w_h \in V_0^h. \tag{3.4}$$

Then, we have the following error estimates (see, e.g., [4, 27]).

Lemma 3.1. *Let* Q_h *and* R_h *denote the* L^2 *projection operator and Ritz projection operator defined above. Then the following estimates hold:*

$$\|v - Q_h v\|_{-1,\Omega} + h\|v - Q_h v\|_{0,\Omega} \leqslant Ch^2\|v\|_{1,\Omega}, \tag{3.5}$$

$$\|v - R_h v\|_{0,\Omega} + h\|v - Q_h v\|_{1,\Omega} \leqslant Ch^2\|v\|_{2,\Omega}. \tag{3.6}$$

Moreover, we have

$$\|v - R_h v\|_{0,\infty,\Omega} \leqslant Ch^{2-\frac{d}{2}}\|v\|_{2,\Omega}. \tag{3.7}$$

Let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a time grid with $k_n = nk$, $n = 1, 2, \cdots, N$, where $k := \frac{T}{N}$. Set $I_n = (t_{n-1}, t_n]$. For $n = 1, 2, \cdots, N$, we construct finite element spaces $V^h \subset C(\bar\Omega)$ associated with the mesh $\mathscr{T}^h$. In what follows we suppose $k = O(h^d)$ for our error analysis.

Note that the weak formulation of the state equation is given by

$$-(y, \partial_t v)_{\Omega_T} + (y, A^* v)_{\Omega_T} = \langle Bu, v\rangle_{\Omega_T} + (y_0, v(\cdot, 0))$$

for every $v \in W(\Omega)$. Then we can define the fully discrete finite element approximation to the state equation as

$$\begin{cases} (\dfrac{y_h^n - y_h^{n-1}}{k}, v_h) + a(y_h^n, v_h) = \langle Bu, v_h\rangle_{I_n} \quad \forall\, v_h \in V_0^h, \\ \qquad\qquad y_h^0 = y_0^h,\ n = 1, \cdots, N, \end{cases} \tag{3.8}$$

where $y_0^h = Q_h y_0(x)$. Here

$$\langle Bu, v_h\rangle_{I_n} = \frac{1}{k}\int_{t_{n-1}}^{t_n} (Bu, v_h)dt = \frac{1}{k}\int_{t_{n-1}}^{t_n} (\sum_{i=1}^m u_i(t) v_h(X_i))dt \quad \forall\, v_h \in V_0^h.$$

For this discrete scheme we have the following stability estimates.

Lemma 3.2. *Assume that $y_h^n \in V_0^h, n = 1, 2, \cdots, N$ are the solutions of the fully discrete scheme (3.8). Let $y_0 \in L^2(\Omega)$ and $u \in L^2(0,T;\mathbb{R}^m)$. Then there exists a constant C such that*

$$\sum_{n=1}^{N} \|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + ka(y_h^N, y_h^N) \leqslant C\big(\|y_0\|_{0,\Omega}^2 + \|u\|_{L^2(0,T;\mathbb{R}^m)}^2\big) \tag{3.9}$$

and

$$\|y_h^N\|_{0,\Omega}^2 + k\sum_{n=1}^{N} a(y_h^n, y_h^n) \leqslant C\big(\|y_0\|_{0,\Omega}^2 + \|u\|_{L^2(0,T;\mathbb{R}^m)}^2\big). \tag{3.10}$$

Proof. Setting $v_h = k(y_h^n - y_h^{n-1})$ in (3.8) we obtain

$$\|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + ka(y_h^n, y_h^n - y_h^{n-1}) = k\langle Bu, y_h^n - y_h^{n-1}\rangle_{I_n}.$$

Furthermore, by Young's inequality we have

$$\begin{aligned}
&\|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + ka(y_h^n, y_h^n - y_h^{n-1})\\
\leqslant\ & \|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + \frac{1}{2}k[a(y_h^n, y_h^n) - a(y_h^{n-1}, y_h^{n-1})]\\
\leqslant\ & \sum_{i=1}^{m}\int_{t_{n-1}}^{t_n} u_i(t)(y_h^n - y_h^{n-1})(X_i)dt\\
\leqslant\ & Ck^{\frac{1}{2}}\|y_h^n - y_h^{n-1}\|_{0,\infty,\Omega}\|u\|_{L^2(t_{n-1},t_n;\mathbb{R}^m)}\\
\leqslant\ & \frac{1}{2}\|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + C\|u\|_{L^2(t_{n-1},t_n;\mathbb{R}^m)}^2.
\end{aligned}$$

Here we used the inverse estimate

$$k^{\frac{1}{2}}\|y_h^n - y_h^{n-1}\|_{0,\infty,\Omega} \leqslant Ck^{\frac{1}{2}}h^{-\frac{d}{2}}\|y_h^n - y_h^{n-1}\|_{0,\Omega} \leqslant C\|y_h^n - y_h^{n-1}\|_{0,\Omega},$$

where the latter estimate follows from our assumption $k = O(h^d)$. Summation from $n = 1$ to N leads to

$$\begin{aligned}
\sum_{n=1}^{N} \|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + ka(y_h^N, y_h^N) &\leqslant Ck\|y_h^0\|_{1,\Omega}^2 + C\|u\|_{L^2(0,T;\mathbb{R}^m)}^2\\
&\leqslant Ck\|Q_h y_0\|_{1,\Omega}^2 + C\|u\|_{L^2(0,T;\mathbb{R}^m)}^2\\
&\leqslant C\|y_0\|_{0,\Omega}^2 + C\|u\|_{L^2(0,T;\mathbb{R}^m)}^2,
\end{aligned}$$

which implies (3.9).

By choosing $v_h = ky_h^n$ in (3.12) and proceeding similarly as the proof of (3.9) we can obtain the stability estimate (3.10). □

We use the variational discretization approach developed in [15] to discretize our optimal control problem. Now we are in a position to introduce the fully discretized optimal control problem: find $(y_h^n, u_h(t)) \in V_0^h \times U_{ad}$ satisfying

$$\min_{u_h \in U_{ad}} J(y_h, u_h) = \frac{1}{2}k\sum_{n=1}^{N}\|y_h^n - y_d^n\|^2 + \frac{\alpha}{2}\|u_h(t)\|_{L^2(0,T;\mathbb{R}^m)}^2 \tag{3.11}$$

subject to

$$\begin{cases} (\dfrac{y_h^n - y_h^{n-1}}{k}, v_h) + a(y_h^n, v_h) = \langle Bu_h(t), v_h\rangle_{I_n} \quad \forall\, v_h \in V_0^h,\\ \qquad n = 1, 2, \cdots, N, \quad y_h^0 = y_0^h \ \text{ in } \Omega, \end{cases} \tag{3.12}$$

where $u_h(t) = \big(u_{1,h}(t), \cdots, u_{m,h}(t)\big)$ and $y_h^n = y_h^n(u_h)$ denotes the solution to (3.8) with $u = u_h$.

Problem (3.11) admits a unique solution by standard arguments. Moreover, we have the following discrete first order optimality conditions: there exists a unique discrete adjoint state $p_h^{n-1} \in V_0^h$ such that

$$
(3.13) \qquad \begin{cases} (\dfrac{p_h^{n-1}-p_h^n}{k}, w_h) + a(w_h, p_h^{n-1}) = (y_h^n - y_d^n, w_h) \quad \forall\, w_h \in V_0^h, \\ p_h^N = 0, \; n = 0, \cdots, N-1 \end{cases}
$$

and

$$
(3.14) \qquad \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} (\alpha u_{i,h}(t) + p_h^{n-1}(X_i))(\tilde{v}_i(t) - u_{i,h}(t))dt \geqslant 0 \qquad \forall\, \tilde{v}(t) \in U_{ad}.
$$

4. Error analysis of optimal control problems

This section is devoted to the error analysis for the finite element approximation of optimal control problems. In the subsequent analysis we need to introduce the following two auxiliary problems: for $u_h(t) \in U_{ad}$, find $y(u_h) \in L^2(0,T;L^2(\Omega))$ satisfying

$$
(4.1) \qquad \begin{cases} \partial_t y(u_h) + Ay(u_h) = \displaystyle\sum_{i=1}^{m} u_{i,h}(t)\delta_{X_i} \quad \text{in} \;\; \Omega_T, \\ y(u_h)(x,t) = 0 \;\; \text{on} \; \Gamma_T, \\ y(u_h)(x,0) = y_0(x) \;\; \text{in} \; \Omega \end{cases}
$$

and for $Y_h \in V_0^h$, find $p(Y_h) \in L^2(0,T;L^2(\Omega))$ satisfying

$$
(4.2) \qquad \begin{cases} -\partial_t p(Y_h) + A^* p(Y_h) = Y_h - y_d \quad \text{in} \;\; \Omega_T, \\ p(Y_h)(x,t) = 0 \;\; \text{on} \; \Gamma_T, \\ p(Y_h)(x,T) = 0 \;\; \text{in} \; \Omega, \end{cases}
$$

where $Y_h|_{I_n} = y_h^n, n = 1, 2, \cdots, N$. We also set $P_h|_{I_n} = p_h^{n-1}$.

Since $u_h(t) \in U_{ad} \subset L^2(0,T;\mathbb{R}^m)$, one concludes from Theorem 2.1 that problem (4.1) admits a unique solution $y(u_h) \in L^2(0,T;L^2(\Omega))$. Similarly, problem (4.2) admits a unique solution $p(Y_h) \in L^2(0,T;H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$.

Here $y_h^n \in V_0^h$ and $p_h^n \in V_0^h$, $n = 1, \cdots, N$ are the fully discrete finite element approximations of $y(u_h)$ and $p(Y_h)$, respectively. We have

Lemma 4.1. *Let $y(u_h) \in L^2(0,T;L^2(\Omega))$ and $y_h^n \in V_0^h$, $n = 1, \cdots, N$ be the solutions of problem (4.1) and (3.12), respectively. Then we have the following a priori error estimate:*

$$
(4.3) \qquad \|y(u_h) - Y_h\|_{L^2(0,T;L^2(\Omega))} \leqslant C(h^{2-\frac{d}{2}} + k^{\frac{1}{2}}).
$$

Proof. The proof follows the ideas of [13], see also [10]. For the estimate of $\|y(u_h) - Y_h\|_{L^2(0,T;L^2(\Omega))}$ we use a duality argument. Consider the dual problems

$$
(4.4) \qquad \begin{cases} -\partial_t \psi + A^*\psi = f \;\; \text{in} \; \Omega_T, \\ \psi = 0 \;\; \text{in} \; \Gamma_T, \\ \psi(x,T) = 0 \;\; \text{in} \; \Omega. \end{cases}
$$

Then, according to [19] for $f \in L^2(0,T;L^2(\Omega))$ the following stability estimates hold:

$$
(4.5) \qquad \|\psi\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t \psi\|_{L^2(0,T;L^2(\Omega))} \;\; \leqslant \;\; C\|f\|_{L^2(0,T;L^2(\Omega))}
$$

and

$$
(4.6) \qquad \|\psi(0)\|_{1,\Omega} \;\; \leqslant \;\; C\|f\|_{L^2(0,T;L^2(\Omega))}.
$$

Furthermore, we have

$$
\int_0^T \int_\Omega (y(u_h) - Y_h) f dxdt
$$

$$
\begin{aligned}
&= \int_0^T (y(u_h) - Y_h, -\partial_t\psi + A^*\psi)dt \\
&= (y_0(x), \psi(0)) + \sum_{i=1}^m \int_0^T (u_{i,h}(t)\delta_{X_i}, \psi)dt - \int_0^T (Y_h, -\partial_t\psi + A^*\psi)dt \\
&= (y_0(x), \psi(0)) + \sum_{i=1}^m \int_0^T (u_{i,h}(t)\delta_{X_i}, \psi)dt - \sum_{n=1}^N (y_h^n - y_h^{n-1}, \psi^{n-1}) \\
&\quad -(y_0^h, \psi(0)) - \sum_{n=1}^N \int_{I_n} a(y_h^n, \psi)dt \\
&= (y_0(x) - y_0^h, \psi(0)) + \sum_{i=1}^m \int_0^T (u_{i,h}(t)\delta_{X_i}, \psi)dt - \sum_{n=1}^N (y_h^n - y_h^{n-1}, \psi^{n-1}) \\
&\quad - \sum_{n=1}^N \int_{I_n} a(y_h^n, \psi)dt. \qquad (4.7)
\end{aligned}
$$

From (3.12) we have

$$
\sum_{n=1}^N (y_h^n - y_h^{n-1}, v_h) + \sum_{n=1}^N \int_{I_n} a(y_h^n, v_h)dt = k \sum_{i=1}^m \sum_{n=1}^N \langle u_{i,h}(t)\delta_{X_i}, v_h\rangle_{I_n}.
$$

Setting $v_h = \bar{R}_h\psi = \frac{1}{k}\int_{I_n} R_h\psi dt$ leads to

$$
\sum_{n=1}^N (y_h^n - y_h^{n-1}, \bar{R}_h\psi) + \sum_{n=1}^N \int_{I_n} a(y_h^n, \bar{R}_h\psi)dt = k \sum_{i=1}^m \sum_{n=1}^N \langle u_{i,h}(t)\delta_{X_i}, \bar{R}_h\psi\rangle_{I_n}. \qquad (4.8)
$$

Substituting (4.8) into (4.7) gives

$$
\begin{aligned}
&\int_0^T \int_\Omega (y(u_h) - Y_h) f dxdt \\
&= (y_0(x) - y_0^h, \psi(0)) + \sum_{i=1}^m \int_0^T (u_{i,h}(t)\delta_{X_i}, \psi)dt - k\sum_{i=1}^m \sum_{n=1}^N \langle u_{i,h}(t)\delta_{X_i}, \bar{R}_h\psi\rangle_{I_n} \\
&\quad - \sum_{n=1}^N (y_h^n - y_h^{n-1}, \psi^{n-1} - \bar{R}_h\psi) - \sum_{n=1}^N \int_{I_n} a(y_h^n, \psi - \bar{R}_h\psi)dt \\
&= T_1 + T_2 + T_3. \qquad (4.9)
\end{aligned}
$$

Now it remains to investigate the addends of T_1, T_2 and T_3. By choosing $y_0^h = Q_h y_0$, from (3.5) we have

$$
\begin{aligned}
T_1 &\leqslant \|y_0 - y_0^h\|_{-1,\Omega} \|\psi(0)\|_{1,\Omega} \leqslant Ch\|y_0\|_{0,\Omega}\|\psi(0)\|_{1,\Omega} \\
&\leqslant Ch\|y_0\|_{0,\Omega}\|f\|_{L^2(0,T;L^2(\Omega))} \\
&\leqslant Ch^{2-\frac{d}{2}}\|y_0\|_{0,\Omega}\|f\|_{L^2(0,T;L^2(\Omega))}. \qquad (4.10)
\end{aligned}
$$

For T_2 we have

$$
\begin{aligned}
T_2 &= \sum_{i=1}^m \int_0^T (u_{i,h}(t)\delta_{X_i}, \psi)dt - k\sum_{i=1}^m \sum_{n=1}^N \langle u_{i,h}(t)\delta_{X_i}, \bar{R}_h\psi\rangle_{I_n} \\
&= \sum_{i=1}^m \sum_{n=1}^N \int_{I_n} u_{i,h}(t)\psi(X_i, t)dt - \sum_{i=1}^m \sum_{n=1}^N \int_{I_n} u_{i,h}(t)\bar{R}_h\psi(X_i, t)dt
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} u_{i,h}(t)(\psi(X_i,t)dt - R_h\psi(X_i,t))dt \\
&\leqslant \|u_h\|_{L^2(0,T;\mathbb{R}^m)}\|\psi - R_h\psi\|_{L^2(0,T;L^\infty(\Omega))} \\
&\leqslant Ch^{2-\frac{d}{2}}\|u_h\|_{L^2(0,T;\mathbb{R}^m)}\|f\|_{L^2(0,T;L^2(\Omega))} \qquad (4.11)
\end{aligned}
$$

by (4.5) and well known estimates for the Ritz projection. Since $\int_{I_n}(\psi - \bar\psi)dt = 0$, there holds

$$
\begin{aligned}
|T_3| &= \Big|\sum_{n=1}^{N}(y_h^n - y_h^{n-1}, \psi^{n-1} - \bar R_h\psi) + \sum_{n=1}^{N}\int_{I_n} a(y_h^n, \bar\psi - \bar R_h\psi)dt\Big| \\
&= \Big|\sum_{n=1}^{N}(y_h^n - y_h^{n-1}, \psi^{n-1} - \bar R_h\psi) + k\sum_{n=1}^{N} a(y_h^n, \bar\psi - \bar R_h\psi)\Big| \\
&\leqslant \sum_{n=1}^{N}\|y_h^n - y_h^{n-1}\|_{0,\Omega}\|\psi^{n-1} - \bar R_h\psi\|_{0,\Omega} + k\sum_{n=1}^{N} a^{\frac12}(y_h^n, y_h^n)a^{\frac12}(\bar\psi - \bar R_h\psi, \bar\psi - \bar R_h\psi).
\end{aligned}
$$

Standard error estimates yield

$$
\begin{aligned}
\|\psi^{n-1} - \bar R_h\psi\|_{0,\Omega} &\leqslant \|\psi^{n-1} - \bar\psi\|_{0,\Omega} + \|\bar\psi - \bar R_h\psi\|_{0,\Omega} \\
&\leqslant Ck^{\frac12}\|\partial_t\psi\|_{L^2(t_{n-1},t_n;L^2(\Omega))} + Ch^2\|\bar\psi\|_{2,\Omega} \qquad (4.12)
\end{aligned}
$$

and

$$
a(\bar\psi - \bar R_h\psi, \bar\psi - \bar R_h\psi) \leqslant Ch^2\|\bar\psi\|_{2,\Omega}^2. \qquad (4.13)
$$

It is straightforward to show that

$$
\|\bar\psi\|_{2,\Omega} \leqslant k^{-\frac12}\|\psi\|_{L^2(t_{n-1},t_n;H^2(\Omega))}. \qquad (4.14)
$$

Utilizing (4.12)-(4.14), Lemma 3.2 and the stability estimates of ψ we get

$$
\begin{aligned}
|T_3| &\leqslant \Big(\sum_{n=1}^{N}(\|y_h^n - y_h^{n-1}\|_{0,\Omega}^2 + ka(y_h^n, y_h^n))\Big)^{\frac12} \\
&\quad \Big(\sum_{n=1}^{N}(\|\psi^{n-1} - \bar R_h\psi\|_{0,\Omega}^2 + ka(\bar\psi - \bar R_h\psi, \bar\psi - \bar R_h\psi))\Big)^{\frac12} \\
&\leqslant \Big(\sum_{n=1}^{N}((h^4 + kh^2)\|\bar\psi\|_{2,\Omega}^2 + k\|\psi_t\|_{L^2(t_{n-1},t_n;L^2(\Omega))}^2)\Big)^{\frac12}(\|y_0\|_{0,\Omega} + \|u_h\|_{L^2(0,T;\mathbb{R}^m)}) \\
&\leqslant C(h^{2-\frac{d}{2}} + k^{\frac12})(\|y_0\|_{0,\Omega} + \|u_h\|_{L^2(0,T;\mathbb{R}^m)})\|f\|_{L^2(0,T;L^2(\Omega))}. \qquad (4.15)
\end{aligned}
$$

Inserting the estimates for T_1, T_2 and T_3 into (4.9) yields

$$
\begin{aligned}
&\int_0^T\int_\Omega (y(u_h) - Y_h)f\,dxdt \\
&\leqslant C(h^{2-\frac{d}{2}} + k^{\frac12})\big(\|y_0\|_{0,\Omega} + \|u_h\|_{L^2(0,T;\mathbb{R}^m)}\big)\|f\|_{L^2(0,T;L^2(\Omega))}. \qquad (4.16)
\end{aligned}
$$

From the definition of $\|Y_h - y(u_h)\|_{L^2(0,T;L^2(\Omega))}$ we derive

$$
\|Y_h - y(u_h)\|_{L^2(0,T;L^2(\Omega))} \leqslant C(h^{2-\frac{d}{2}} + k^{\frac12})\big(\|y_0\|_{0,\Omega} + \|u_h\|_{L^2(0,T;\mathbb{R}^m)}\big),
$$

which completes the proof. □

Remark 4.2. *In [13] error analysis for parabolic equations with measure data in space are considered, namely, problems with righthand side $\mu = g\omega$ with g and ω are given functions such that $g \in L^2(0,T;\mathcal{C}(\bar\Omega))$ and $\omega \in \mathcal{M}(\bar\Omega)$. An a priori error estimate is obtained there for the fully discrete finite element approximation supposing $g \in L^2(0,T;\mathcal{C}(\bar\Omega)) \cap H^1(0,T;L^\infty(\Omega))$. So we would not expect any convergence order for the error between the solution $y(u)$ of problem (2.3) and the* *solution $Y_h(u)$ of the auxiliary problem (3.12) in the present situation, which is usually exploited in the error analysis for optimal control problems. Fortunately, due to the piecewise constant property of $u_h(t)$ on each time interval I_n we can get error estimates for $\|Y_h - y(u_h)\|_{L^2(0,T;L^2(\Omega))}$ instead.*

Now we will derive a priori error estimates between the solutions of discretized adjoint state equation and the intermediate adjoint state equation (4.2).

Lemma 4.3. *Assume that $p(Y_h) \in L^2(0,T;H^2(\Omega)\cap H_0^1(\Omega))\cap H^1(0,T;L^2(\Omega))$ and $p_h^{n-1} \in V_0^h, n = 1,\cdots,N$ are the solutions of problem (4.2) and (3.13), respectively, then we have*

$$\|p(Y_h) - P_h\|_{L^2(0,T;L^2(\Omega))} \leqslant C(h^2 + k). \tag{4.17}$$

Proof. Note that P_h is the fully discrete finite element approximation of $p(Y_h)$, then it can be proved by standard argument (see [10]) that

$$\begin{aligned}
\|p(Y_h) - P_h\|_{L^2(0,T;L^2(\Omega))} &\leqslant C(h^2 + k)(\|p(Y_h)\|_{L^2(0,T;H^2(\Omega))} + \|p(Y_h)\|_{H^1(0,T;L^2(\Omega))}) \\
&\leqslant C(h^2 + k)\|Y_h - y_d\|_{L^2(0,T;L^2(\Omega))} \\
&\leqslant C(h^2 + k)(\|y_0\|_{0,\Omega} + \|u_h\|_{L^2(0,T;\mathbb{R}^m)} + \|y_d\|_{L^2(0,T;L^2(\Omega))}).
\end{aligned}$$

□

Now we are ready to formulate the main result of this paper, namely, the error estimates between the solutions of the continuous and discretized optimal control problems.

Theorem 4.4. *Let $(y, p, u) \in L^2(0,T;L^2(\Omega)) \times L^2(0,T;H^2(\Omega)\cap H_0^1(\Omega))\cap H^1(0,T;L^2(\Omega)) \times U_{ad}$ and $(y_h^n, p_h^{n-1}, u_h) \in V_0^h\times V_0^h\times U_{ad}$ be the solutions of problems (2.6) and (3.11)-(3.12), respectively. Then there exists a positive constant C such that*

$$\begin{aligned}
&\sqrt{\alpha}\|u - u_h\|_{L^2(0,T;\mathbb{R}^m)} + \|y - Y_h\|_{L^2(0,T;L^2(\Omega))} + \|p - P_h\|_{L^2(0,T;L^2(\Omega))} \\
&\leqslant C(h^{2-\frac{d}{2}} + k^{\frac{1}{2}}). \qquad (4.18)
\end{aligned}$$

Proof. It follows from the continuous and discrete optimality conditions that

$$\sum_{i=1}^{m}\int_0^T (\alpha u_i(t) + p(X_i,t))(\tilde{v}_i(t) - u_i(t))dt \geqslant 0 \quad \forall\ \tilde{v}(t) \in U_{ad} \tag{4.19}$$

and

$$\sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} (\alpha u_{i,h}(t) + p_h^{n-1}(X_i))(\tilde{v}_i(t) - u_{i,h}(t))dt \geqslant 0 \qquad \forall\ \tilde{v}(t) \in U_{ad}. \tag{4.20}$$

Choosing $\tilde{v}_i(t) = u_{i,h}(t)$ in (4.19) as well as $\tilde{v}_i(t) = u_i(t)$ in (4.20), and adding the resulting inequalities leads to

$$\begin{aligned}
\alpha\|u - u_h\|^2_{L^2(0,T;\mathbb{R}^m)} &= \alpha\sum_{i=1}^{m}\int_0^T (u_i(t) - u_{i,h}(t))^2dt \\
&\leqslant \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} (p_h^{n-1}(X_i) - p(X_i,t))(u_i(t) - u_{i,h}(t))dt. \qquad (4.21)
\end{aligned}$$

Using the auxiliary variable $p(Y_h)$ we can rewrite (4.21) as follows

$$\begin{aligned}
\alpha\|u - u_h\|^2_{L^2(0,T;\mathbb{R}^m)} &\leqslant \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} (p_h^{n-1}(X_i) - p(Y_h)(X_i,t))(u_i(t) - u_{i,h}(t))dt \qquad (4.22)\\
&\quad+\sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n} (p(Y_h)(X_i,t) - p(X_i,t))(u_i(t) - u_{i,h}(t))dt.
\end{aligned}$$

Following (2.1) and (4.1) we have

$$-(y - y(u_h), \partial_t v)_{\Omega_T} + (y - y(u_h), A^*v)_{\Omega_T} = \langle B(u - u_h), v\rangle_{\Omega_T}. \tag{4.23}$$

Setting $v = p(Y_h) - p$ in (4.23) yields

$$\begin{aligned} & \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n}(p(Y_h)(X_i,t)-p(X_i,t))(u_i(t)-u_{i,h}(t))dt \\ = & \ \langle (B(u-u_h), p(Y_h)-p\rangle_{\Omega_T} \\ = & -\int_0^T [((y-y(u_h),\partial_t(p(Y_h)-p)) + (y-y(u_h), A^*(p(Y_h)-p))]dt \\ = & \int_0^T (Y_h-y, y-y(u_h))dt \\ = & -\int_0^T \|Y_h-y\|_{0,\Omega}^2 dt + \int_0^T (Y_h-y, Y_h-y(u_h))dt \\ \leqslant & -\frac{1}{2}\|Y_h-y\|^2_{L^2(0,T;L^2(\Omega))} + \frac{1}{2}\|Y_h-y(u_h)\|^2_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

Using Young's inequality gives

$$\begin{aligned} & \sum_{i=1}^{m}\sum_{n=1}^{N}\int_{I_n}(p_h^{n-1}(X_i)-p(Y_h)(X_i,t))(u_i(t)-u_{i,h}(t))dt \\ \leqslant & \ \frac{1}{2}\alpha\|u-u_h\|^2_{L^2(0,T;\mathbb{R}^m)} + C\|P_h-p(Y_h)\|^2_{L^2(0,T;L^\infty(\Omega))}. \end{aligned} \tag{4.24}$$

Therefore, we obtain

$$\begin{aligned} & \alpha\|u-u_h\|^2_{L^2(0,T;\mathbb{R}^m)} + \|y-Y_h\|^2_{L^2(0,T;L^2(\Omega))} \\ \leqslant & \ C\|P_h-p(Y_h)\|^2_{L^2(0,T;L^\infty(\Omega))} + C\|Y_h-y(u_h)\|^2_{L^2(0,T;L^2(\Omega))}. \end{aligned} \tag{4.25}$$

It remains to estimate $\|P_h-p(Y_h)\|_{L^2(0,T;L^\infty(\Omega))}$ and $\|Y_h-y(u_h)\|_{L^2(0,T;L^2(\Omega))}$.

Firstly, by the inverse inequality and the estimate of Ritz projection we deduce

$$\begin{aligned} & \|P_h-p(Y_h)\|_{L^2(0,T;L^\infty(\Omega))} \\ \leqslant & \ \|p(Y_h)-R_hp(Y_h)\|_{L^2(0,T;L^\infty(\Omega))} + \|P_h-R_hp(Y_h)\|_{L^2(0,T;L^\infty(\Omega))} \\ \leqslant & \ Ch^{2-\frac{d}{2}}\|p(Y_h)\|_{L^2(0,T;H^2(\Omega))} + Ch^{-\frac{d}{2}}\|P_h-R_hp(Y_h)\|_{L^2(0,T;L^2(\Omega))} \\ \leqslant & \ Ch^{2-\frac{d}{2}}\|p(Y_h)\|_{L^2(0,T;H^2(\Omega))} + Ch^{-\frac{d}{2}}(\|P_h-p(Y_h)\|_{L^2(0,T;L^2(\Omega))} \\ & +\|p(Y_h)-R_hp(Y_h)\|_{L^2(0,T;L^2(\Omega))}) \\ \leqslant & \ Ch^{2-\frac{d}{2}}\|p(Y_h)\|_{L^2(0,T;H^2(\Omega))} + Ch^{-\frac{d}{2}}\|P_h-p(Y_h)\|_{L^2(0,T;L^2(\Omega))}. \end{aligned} \tag{4.26}$$

Thus, combining (4.26) and (4.17) leads to

$$\begin{aligned} \|P_h-p(Y_h)\|_{L^2(0,T;L^\infty(\Omega))} & \leqslant Ch^{2-\frac{d}{2}} + Ch^{-\frac{d}{2}}(h^2+k), \\ & \leqslant Ch^{2-\frac{d}{2}} + k^{\frac{1}{2}}, \end{aligned} \tag{4.27}$$

where the assumption $k = O(h^d)$ is used.

Combining (4.25), (4.27) and (4.3) we get

$$\begin{aligned} & \alpha\|u-u_h\|^2_{L^2(0,T;\mathbb{R}^m)} + \|y-Y_h\|^2_{L^2(0,T;L^2(\Omega))} \\ \leqslant & \ C(h^{4-d}+k). \end{aligned} \tag{4.28}$$

Using (2.8) and (4.2) we obtain

$$-(\partial_t(p-p(Y_h)), v)_{\Omega_T} + a(v, p-p(Y_h))_{\Omega_T} = (y-Y_h, v)_{\Omega_T}. \tag{4.29}$$

By the standard stability estimate we obtain

$$\|p-p(Y_h)\|_{L^2(0,T;L^2(\Omega))} \leqslant C\|y-Y_h\|_{L^2(0,T;L^2(\Omega))}. \tag{4.30}$$

It follows from (4.27), (4.28), (4.30) and triangle inequality that

$$\begin{aligned} \|p-P_h\|_{L^2(0,T;L^2(\Omega))} &\leqslant \|p-p(Y_h)\|_{L^2(0,T;L^2(\Omega))}+\|p(Y_h)-P_h\|_{L^2(0,T;L^2(\Omega))} \\ &\leqslant C(h^{2-\frac{d}{2}}+k^{\frac{1}{2}}). \end{aligned} \tag{4.31}$$

Thus, (4.28) and (4.31) complete the proof of the theorem. □

Remark 4.5. *From the proof of Theorem 4.4 we have the following error representation:*

$$\begin{aligned} &\sqrt{\alpha}\|u-u_h\|_{L^2(0,T;\mathbb{R}^m)}+\|y-Y_h\|_{L^2(0,T;L^2(\Omega))}+\|p-P_h\|_{L^2(0,T;L^2(\Omega))} \\ \leqslant\ & C(\|P_h-p(Y_h)\|_{L^2(0,T;L^2(\Omega))}+\|P_h-p(Y_h)\|_{L^2(0,T;L^\infty(\Omega))}+\|Y_h-y(u_h)\|_{L^2(0,T;L^2(\Omega))}). \end{aligned}$$

Since $Y_h\in L^\infty(\Omega_T)$, if we assume that $y_d\in L^\infty(\Omega_T)$ we can conclude that $p(Y_h)\in W_s^{2,1}(\Omega_T)$ for all $s<\infty$, where $W_s^{2,1}(\Omega_T)$ is defined as

$$W_s^{2,1}(\Omega_T):=\{y\in L^s(0,T;W^{2,s}(\Omega)),\quad y_t\in L^s(0,T;L^s(\Omega))\}.$$

Thus, the error estimate $\|P_h-p(Y_h)\|_{L^2(0,T;L^\infty(\Omega))}$ could be improved by using the techniques of [14]. However, due to the reduced convergence order of $\|Y_h-y(u_h)\|_{L^2(0,T;L^2(\Omega))}$, the whole error estimate cannot be improved.

Remark 4.6. *Here we only derive a priori error estimates for the fully discrete finite element approximation of control problems with control constraints, it seems that the results cannot be improved for unconstrained case, since in both cases the state equation exhibits low regularity, and thus the estimates for state equation approximation dominate in the error representation.*

5. Extensions

In the above sections we investigated the fully discrete finite element approximation and corresponding error analysis for optimal control of parabolic equation with pointwise control. In many practical applications, especially in the environment science models, the governing equations include convection. In this section we will present the extension to the pointwise control problem with transient convection diffusion equations.

The governing equation can be extended to the following transient convection diffusion problems

$$\begin{cases} y_t+\tilde{A}y=\sum_{i=1}^m u_i(t)\delta_{X_i} & \text{in } \Omega_T, \\ y(x,t)=0 & \text{on } \Gamma_T, \\ y(x,0)=y_0(x) & \text{in } \Omega, \end{cases} \tag{5.1}$$

where the operator $\tilde{A}$ takes the following form

$$\tilde{A}y=-\sum_{i,j=1}^d \partial_{x_j}(a_{ij}\partial_{x_i}y)+\sum_{i=1}^d \beta_i\partial_{x_i}y+\tilde{a}_0y,$$

$\boldsymbol{\beta}=(\beta_1,\beta_2,\cdots,\beta_d)^T$ denotes the velocity field, $\tilde{a}_0>$ denotes the reaction coefficient. We assume that $\tilde{a}_0-\frac{1}{2}\sum_{i=1}^d\partial_{x_i}\beta_i>\gamma_0>0$. Let $\tilde{A}^*$ be the adjoint operator of $\tilde{A}$, which can be characterized by

$$\tilde{A}^*y=-\sum_{i,j=1}^d \partial_{x_j}(a_{ji}\partial_{x_i}y)-\sum_{i=1}^d \partial_{x_i}(\beta_iy)+\tilde{a}_0y.$$

Suppose that the location X_i are given, one needs to determine the magnitude $u_i(t)$ at X_i through some observations y_d, which is called data assimilation problems in the context of inverse problems. We can reformulate the inverse problem through an output least square approach and Tikhonov regularization to the following pointwise control problem

$$(5.2)\qquad \begin{cases} \min\limits_{u\in U_{ad}} \ J(y,u) = \dfrac{1}{2}\|y-y_d\|^2_{L^2(\Omega_T)} + \dfrac{\alpha}{2}\|u\|^2_{L^2(0,T;\mathbb{R}^m)} \\ \qquad\qquad \text{subject to } \ (5.1) \end{cases}$$

with regularization parameter $\alpha > 0$. By using similar arguments as in Section 2 we can prove the existence of weak solutions to the state equation by transposition techniques, and thus the existence and uniqueness of a solution for the optimal control problem. The fully discrete finite element approximation as well as error estimates can be derived similarly.

Remark 5.1. *In case that the state equation of optimal control problem (5.2) is diffusion dominated we can use the discrete scheme proposed for heat equation in Section 3 for its numerical approximation, and the corresponding error estimates are also valid to this case. But if the state equation is convection dominated we have to adopt other stable numerical methods, such as stabilized finite element methods. A more natural choice for such kind of problems may be the characteristic finite element method proposed in* [8]*. Error analysis for this scheme in the case of distributed controls is presented in* [11].

In environmental science models, such as air pollution and waste water treatment problems, the optimal control problem (5.2) is often accompanied with state constraints in several local zones. In the following let's consider a mathematical model to illustrate this situation. It reads

$$(5.3)\qquad \min_{u\in U_{ad}} \quad J(y,u) = \tfrac{1}{2}\int_0^T\int_\Omega (y-y_d)^2 dxdt + \tfrac{1}{2}\sum_{i=1}^m \int_0^T u_i(t)^2\omega_i(t)dt$$

subject to

$$(5.4)\qquad \begin{cases} y_t + \tilde{A}y = \displaystyle\sum_{i=1}^m E_i(t)(1-u_i(t))\delta_{X_i} & \text{in } \Omega_T, \\ y(x,t) = 0 & \text{on } \Gamma_T, \\ y(x,0) = y_0(x) & \text{in } \Omega \end{cases}$$

and

$$(5.5)\qquad y(x,t) \leqslant \sigma_j \ \text{ in } \ D_j \times [0,T], \ j = 1,2,\cdots,N_d,$$

where $\omega_i(t)$ are the weight functions measuring the cost of imposing control, and $E_i(t)$ denote the maximum emissions of pollutants at time t. The admissible control set is $U_{ad} := \{u \in L^2(0,T;\mathbb{R}^m) : 0 \leqslant u_i(t) \leqslant 1, \ \ i = 1,2,\cdots,m \ \ a.e. \ in \ \ (0,T)\}$. $D_i \subset\subset \Omega$ denote the observation areas, which are supposed to be away from the pollution sources X_i, σ_j are given tolerances and N_d is a positive integer, see Fig. 5.1.

In this kind of problems y denotes the concentration of pollutant resulting from the emission source located at X_i (factory or waste water plant). One needs to control the emission of pollutant in X_i such that the pollutant concentrations in observation area D_i (school, hospital or residence) are below the tolerance level σ_i. An example of this kind of problem is sketched in Figure 5.1.

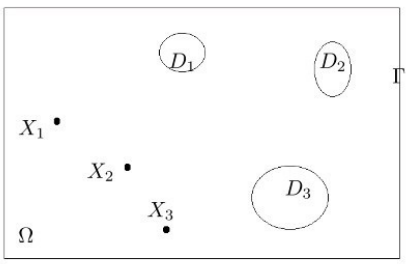


Figure 5.1

Note that the state constraints (5.5) can be treated by penalty method, i.e., we can penalize the state constraints in the objective functional through least square approach, then problems (5.3)-(5.5) can be analyzed in the framework of (5.2) after minor modifications.

Go back to the original state constrained problem, the existence of weak solution of the state equation (5.1) can be obtained by transposition techniques similar to Theorem 2.1. We can also prove the existence and uniqueness of solutions for the optimal control problems by standard arguments.

To derive the first order optimality condition we assume the following Slater condition: there exists a feasible control $\tilde{u} \in U_{ad}$ such that

$$\tilde{y}(x,t) < \sigma_j \quad \text{in} \quad D_j \times [0,T], \ j = 1,2,\cdots,N_d, \tag{5.6}$$

where $\tilde{y}$ is the state associated with $\tilde{u}$. Then we are able to derive the following first order necessary conditions:

Theorem 5.2. *Suppose that $u \in U_{ad}$ is the solution to problem (5.3) with associated state $y \in L^2(0,T;L^2(\Omega))$. Let $\mathcal{M}(\bar{\Omega}_T)$ be the space of regular Borel measures on $\bar{\Omega}_T$. Then there exists an adjoint state $p \in L^q(0,T;W^{1,\sigma}(\Omega))$ for all $q,\sigma \in [1,2)$ with $\frac{2}{q} + \frac{d}{\sigma} > d+1$, and a Lagrange multiplier $\mu \in \mathcal{M}(\bar{\Omega}_T)$ satisfying*

$$\begin{cases} -p_t + \tilde{A}^* p = y - y_d + \mu|_{\Omega\times(0,T)} & \text{in } \Omega_T, \\ p(x,t) = 0 & \text{on } \Gamma_T, \\ p(x,T) = \mu|_{\Omega\times\{T\}} & \text{in } \Omega \end{cases} \tag{5.7}$$

in the sense of distribution, and the following relation holds

$$\sum_{i=1}^{m} \int_0^T (u_i(t)\omega_i(t) - E_i(t)p(X_i,t))(\tilde{v}_i - u_i)dt \geqslant 0 \quad \forall\ \tilde{v} \in U_{ad}. \tag{5.8}$$

Proof. The proof can be found in, e.g., [2], [14] and [23]. □

Despite the low global regularity of state y, following the ideas of [23] we can prove that the state y is continuous in $\bar{D}_j \times [0,T]$, $j = 1,\cdots,N_d$, and thus the state constraints (5.5) are well defined. Note that the state constraints (5.5) are only imposed in subdomain D_j, thus the multiplier μ associated with state constraints is a Borel measure with support in $\cup_{j=1}^{N_d} \bar{D}_j \times [0,T]$. This property implies the continuity of the adjoint state p in the subdomain containing the points X_i, and thus the optimality condition (5.8) which is associated with point value of p at X_i is also well defined.

Remark 5.3. *The presence of state constraints makes the optimal control problem (5.3) more complicated than (5.2), not only in choosing a suitable numerical approximation scheme, but also in deriving the corresponding error analysis, which we postpone to future work.*

6. Numerical Experiments

In this section we will carry out some numerical experiments to confirm our theoretical findings. In the first two examples we consider the following parabolic optimal control problem with pointwise control:

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

subject to

$$\begin{cases} y_t - \Delta y = f + u(t)\cdot\delta_X & \text{in } \Omega_T, \\ y(x,t) = 0 & \text{on } \Gamma_T, \\ y(x,0) = 0 & \text{in } \Omega, \end{cases} \tag{6.1}$$

where we choose $(Bu)(x,t) = u(t)\cdot\delta_X$, so that $m = 1$ in (2.1), δ_X is the Dirac Delta function which will be specified in the following examples. For ease of constructing examples we may admit some additional regular parts f to appear in the righthand side.

For the computation the software package AFEPack ([17]) is used. We use a preconditioned projection algorithm (see, Section 8.2 in [21], for more details) to solve the discretized optimization problems. In the following numerical examples, we define an error functional E and illustrate its experimental order of convergence by

$$\text{rate} = \frac{\log E(h_1) - \log E(h_2)}{\log h_1 - \log h_2},$$

where h denotes the mesh size of the triangulation or the time step k.

Example 6.1. *The first example is a modification of the example in* [13]. *Let* $\Omega_T = B(0,1) \times [0,1]$, *where* $B(0,1)$ *is the unit circle centered at zero with radius 1. In this example the control is unconstrained, i.e.,* $U_{ad} = L^2(0,T)$. *We take the exact solutions as*

$$y(x,t) = -\frac{1}{2\pi}\log|x| \cdot t(1-t), \quad u(t) = t(1-t),$$

and

$$p(x,t) = -\cos(\frac{\pi}{2}|x|^2) \cdot t(1-t).$$

Then after simple calculation we have

$$\begin{aligned} f(x,t) &= -\frac{1}{2\pi}\log|x| \cdot (1-2t), \\ y_d(x,t) &= -\frac{1}{2\pi}\log|x| \cdot t(1-t) - \cos(\frac{\pi}{2}|x|^2) \cdot (1-2t) \\ &\quad +(2\pi\sin(\frac{\pi}{2}|x|^2) + \pi^2|x|^2\cos(\frac{\pi}{2}|x|^2)) \cdot t(1-t) \end{aligned}$$

and thus $\delta_X = \delta_0$, *where* δ_0 *is the Dirac function at* $x = (0,0)$.

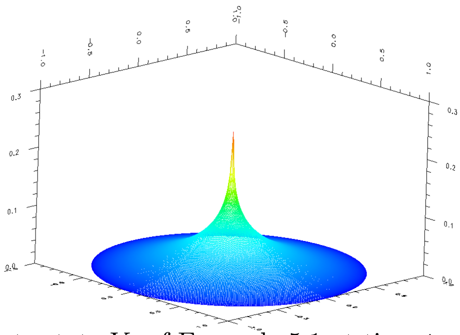


Figure 6.1. The discrete state Y_h of Example 5.1 at time $t = 0.5$ with 16641 Dofs.

At first, we set $k = O(h^2)$, where k is the time triangulation step and h is the mesh size of the spatial triangulation. We see from Table 6.1 that the convergence order for control u and adjoint state p is 2, while the convergence order for state y is only 1, which is consistent with the results presented in [13] where finite element method for parabolic equation with measure data in space is studied. Figure 6.1 presents discrete state Y_h at time $t = 0.5$ with 16641 Dofs, while Figure 6.2 presents exact and discrete control with time step $N = 128$ and Dof $= 16641$.

Table 6.1. Error of control u, state y and adjoint state p for Example 6.1 with respect to space and time.

Dof	N	$\|u-u_h\|_{L^2(0,T)}$	rate	$\|y-Y_h\|_{L^2(\Omega_T)}$	rate	$\|p-P_h\|_{L^2(\Omega_T)}$	rate
25	10	0.033925322659	\	0.004994662884	\	0.056119454603	\
81	40	0.008714930485	1.9608	0.001986026439	1.3305	0.014588041781	1.9437
289	160	0.002187056856	1.9945	0.000917225519	1.1145	0.003679696210	1.9871
1089	640	0.000546232163	2.0014	0.000449061987	1.0304	0.000922243592	1.9964
4225	2560	0.000136261372	2.0031	0.000223315698	1.0078	0.000230721934	1.9990

To validate the estimates developed in the previous section precisely, we show also the convergence order by separating the discretization errors. To investigate the convergence order with respect to space discretization we fixed the time discretization with $N = 4096$, while space discretization is fixed with $Dof = 16641$ to investigate convergence order with respect to time. We can see from Table 6.2 that the convergence order w.r.t space discretization is 2 for the control and the adjoint state, and is 1 for state. Form Table 6.3 we deduce that the convergence order w.r.t time discretization is 1 for both control, state and adjoint state, which is better than our predicted result of order $k^{\frac{1}{2}}$.

Table 6.2. Error of control u, state y and adjoint state p for Example 6.1 with respect to space with fixed time step N=4096.

Dof	$\|u-u_h\|_{L^2(0,T)}$	rate	$\|y-Y_h\|_{L^2(\Omega_T)}$	rate	$\|p-P_h\|_{L^2(\Omega_T)}$	rate
25	0.013548983759	\	0.004782786173	\	0.036175778204	\
81	0.003176540490	2.0927	0.001992216424	1.2635	0.009348190813	1.9523
289	0.000765919014	2.0522	0.000925404659	1.1062	0.002351214586	1.9913
1089	0.000198998000	1.9444	0.000450454865	1.0387	0.000592324225	1.9889
4225	0.000091303265	1.1240	0.000223313048	1.0123	0.000181355022	1.7076

Table 6.3. Error of control u, state y and adjoint state p for Example 6.1 with respect to time with fixed Dof= 16641.

N	$\|u-u_h\|_{L^2(0,T)}$	rate	$\|y-Y_h\|_{L^2(\Omega_T)}$	rate	$\|p-P_h\|_{L^2(\Omega_T)}$	rate
4	0.067067104186	\	0.009378517639	\	0.120265193582	\
8	0.037073940199	0.8552	0.005357366365	0.8078	0.060640796847	0.9879
16	0.019651447523	0.9158	0.002873332814	0.8988	0.030425177713	0.9950
32	0.010144314919	0.9540	0.001492047606	0.9454	0.015253525816	0.9961
64	0.005157965519	0.9758	0.000764572991	0.9646	0.007639859295	0.9975
128	0.002601168197	0.9876	0.000395845333	0.9497	0.003822731668	0.9989
256	0.001306059354	0.9939	0.000219148203	0.8530	0.001911096709	1.0002

Example 6.2. *The second example is similar to Example 6.1 but with control constraints. Let $\Omega_T = B(0,1)\times[0,1]$. We set $a=-1$ and $b=-0.5$ and take the exact solutions as*

$$y(x,t) = -\frac{1}{2\pi}\log|x|\cdot t(exp(t)-exp(T)), \quad u(t) = P_{U_{ad}}(exp(t)-exp(T)),$$

and

$$p(x,t) = -\cos(\frac{\pi}{2}|x|^2)\cdot(exp(t)-exp(T)).$$

Then after simple calculation we have

$$\begin{aligned} f(x,t) &= -\frac{1}{2\pi}\log|x|\cdot(t\cdot exp(t)+exp(t)-exp(T)) \\ &\quad +(t(exp(t)-exp(T)) - P_{U_{ad}}(exp(t)-exp(T)))\cdot\delta_0, \\ y_d(x,t) &= -\frac{1}{2\pi}\log|x|\cdot(exp(t)-exp(T)) - \cos(\frac{\pi}{2}|x|^2)\cdot exp(t) \\ &\quad +(2\pi\sin(\frac{\pi}{2}|x|^2)+\pi^2|x|^2\cos(\frac{\pi}{2}|x|^2))\cdot(exp(t)-exp(T)) \end{aligned}$$

and thus $\delta_X=\delta_0$, where δ_0 is the Dirac function at $x=(0,0)$.

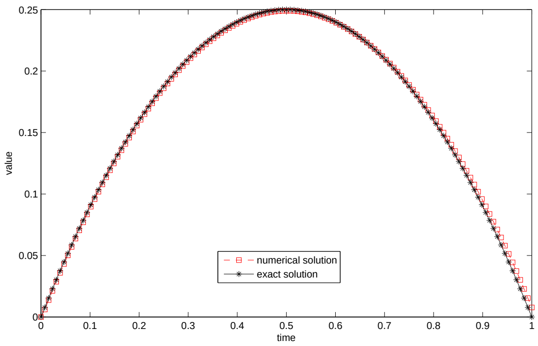


Figure 6.2. The exact solution u versus discrete solution u_h of Example 6.1 with time step $N = 128$ and Dof $= 16641$.

Similar to Example 6.1 we investigate the convergence order with respect to space discretization with fixed time discretization $N = 4096$, while space discretization is fixed with $Dof = 16641$ to investigate convergence order with respect to time. We can see from Table 6.4 that the convergence order w.r.t space discretization is 2 for control and adjoint state, and is 1 for state. Form Table 6.5 we know that the convergence order w.r.t time discretization is 1 for both control, state and adjoint state, which is better than our predicted result of order $k^{\frac{1}{2}}$. The exact solution u versus discrete solution U_h with time step $N = 128$ and Dof $= 16641$ is shown in Figure 6.3.

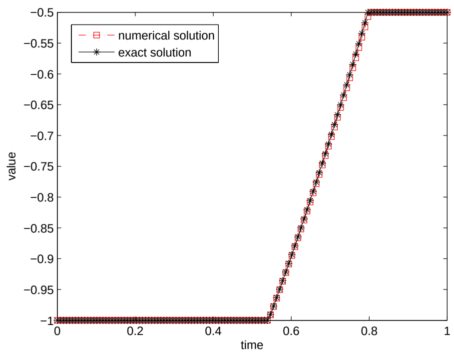


Figure 6.3. The exact solution u versus discrete solution u_h of Example 6.2 with time step $N = 128$ and Dof $= 16641$.

Table 6.4. Error of control u, state y and adjoint state p for Example 6.2 with respect to space with fixed time step N=4096.

Dof	$\|u - u_h\|_{L^2(0,T)}$	rate	$\|y - Y_h\|_{L^2(\Omega_T)}$	rate	$\|p - P_h\|_{L^2(\Omega_T)}$	rate
25	0.029747159840	\	0.009464085679	\	0.216230229539	\
81	0.007338344473	2.0192	0.004159874555	1.1859	0.055650361525	1.9581
289	0.001833890291	2.0005	0.001975677926	1.0742	0.014110263603	1.9796
1089	0.000484266593	1.9210	0.000970029553	1.0262	0.003677734951	1.9399
4225	0.000155122674	1.6424	0.000482138061	1.0086	0.001114196704	1.7228

Table 6.5. Error of control u, state y and adjoint state p for Example 6.2 with respect to time with fixed Dof= 16641.

N	$\|u-u_h\|_{L^2(0,T)}$	rate	$\|y-Y_h\|_{L^2(\Omega_T)}$	rate	$\|p-P_h\|_{L^2(\Omega_T)}$	rate
4	0.056398954132	\	0.012280726137	\	0.430787997471	\
8	0.025966660812	1.1190	0.005957590608	1.0436	0.218313923232	0.9806
16	0.012875983101	1.0120	0.002798312653	1.0902	0.110397317872	0.9837
32	0.006993652675	0.8806	0.001430527113	0.9680	0.055628886472	0.9888
64	0.003798931334	0.8805	0.000725522163	0.9795	0.027973849440	0.9918
128	0.001876036589	1.0179	0.000418171820	0.7949	0.014059665905	0.9925

Example 6.3. *In the third example we consider an optimal control problems with convection term and multiple Dirac points:*

$$\min_{u\in U_{ad}} J(y,u) = \frac{\alpha}{2}\int_0^T\int_\Omega (y-y_d)^2 dxdt + \frac{1}{2}\sum_{i=1}^m \int_0^T u_i(t)^2\omega_i(t)dt$$

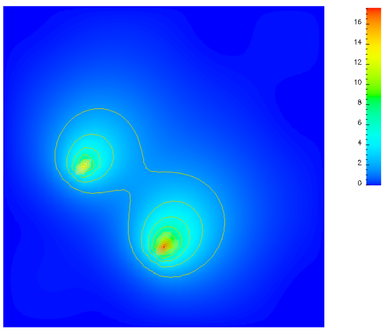
subject to

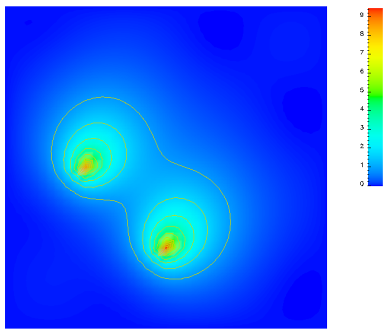
$$\left\{\begin{aligned} y_t - \nabla\cdot(a\nabla y) + \vec{b}\cdot\nabla y + cy = f + \sum_{i=1}^{2} E_i(t)(1-u_i(t))\cdot\delta_{X_i} \quad & in\ \Omega_T, \\ y(x,t) = 0 \quad & on\ \Gamma_T, \\ y(x,0) = 0 \quad & in\ \Omega. \end{aligned}\right. \tag{6.2}$$

Let $\Omega_T = [-4,4]^2\times[0,0.8]$. *We set* $X_1 = (-1,0)$, $X_2 = (0,-1)$. *We also set* $a = 1$, $\vec{b} = (2,3)$, $c = 1$, $\alpha = 1$, $\omega_1 = 40$ *and* $\omega_2 = 40$. $E_1(t)$ *and* $E_2(t)$ *are assumed to be 25 and 35, respectively. We take* f *and the desired state* y_d *as*

$$\begin{aligned} f(x,t) &= \pi t\sin(\pi x_1)\sin(\pi x_2), \\ y_d(x,t) &= -0.05\pi(\log\frac{|x-X_1|}{4} + \log\frac{|x-X_2|}{4}). \end{aligned}$$

This example can be viewed as a model for air pollution control problem. The state y represents the concentration of pollutant, while u denotes the control action. E_1^{max} and E_2^{max} are the maximum emission at points X_1 and X_2 where pollution sources are located. f represents the uncontrollable emissions of some specific polluting sources. At first we consider the case without control, which means only the state equation (6.2) with maximum emissions are solved. Then we solve the optimization problem with control action and desired concentration y_d.





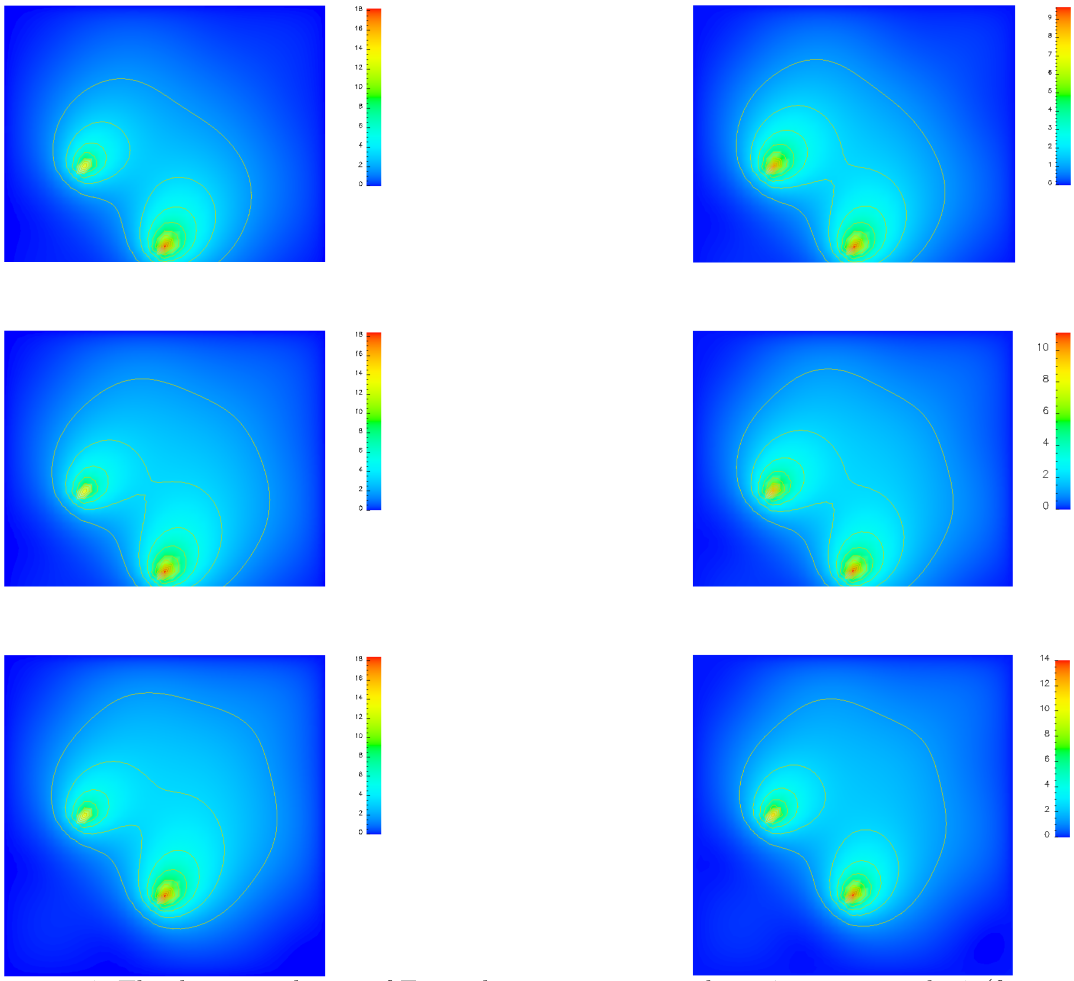


Figure 6.4. The discrete solution of Example 6.3 at time step $k = 64, 128, 192$ and 240(from top to bottom). The left subplots show the solutions without control, while the right subplots show the solutions with control action.

The discretization is based on piecewise linear finite element space with 1681 Dofs and dG(0) in time discretization with 256 steps. The discrete solutions y_h for the uncontrolled and the controlled case on different time step are presented in Figure 6.4, we also present the profiles of discrete controls in Figure 6.5. From Figure 6.4 we can observe that the control action can significantly reduce the concentration of pollutant.

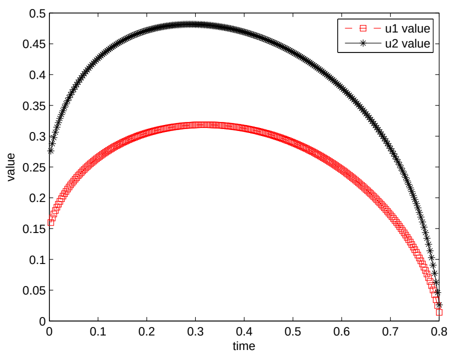


Figure 6.5. The discrete solution $u_{1,h}$ and $u_{2,h}$ of Example 6.3.

Acknowledgements

The first author would like to thank the support of Alexander von Humboldt Foundation during the stay in University of Hamburg, Germany. The second and the third authors gratefully acknowledge the support of the DFG Priority Program 1253 entitled Optimization with Partial Differential Equations. The first and the third authors are also very grateful to the Department of Mathematics, University of Hamburg for the hospitality and support.

References

[1] E. Casas, L^2 estimates for the finite element method for the Dirichlet problem with singular data, Numer. Math., 47 (1985), pp. 627-632.

[2] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optim., 35 (1997), pp. 1297-1327.

[3] I. Chryssoverghi, Approximate methods for optimal pointwise control of parabolic systems, Systems Control Lett., 1 (1981/82), no. 3, pp. 216-229.

[4] P. G. Ciarlet, The finite element methods for elliptic problems, North-Holland, Amsterdam, 1978.

[5] J. C. De Los Reyes, P. Merino, J. Rehberg and F. Tröltzsch, Optimality conditions for state-constrained PDE control problems with time-dependent controls, Control and Cybernetics, 37 (2008), pp. 5-38.

[6] E. J. Dean and P. Gubernatis, Pointwise control of Burgers' equation - a numerical approach, Computers Math. Applic., 22 (1991), pp. 93-100.

[7] K. Deckelnick, M. Hinze, Variational discretization of parabolic control problems in the presence of pointwise state constraints, J. Comput. Math., 29 (2011), pp. 1-16.

[8] J. Douglas Jr. and T. F. Russell, Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, SIAM J. Numer. Anal., 19 (1982), pp. 871-885.

[9] J. Droniou, J. P. Raymond, Optimal pointwise control of semilinear parabolic equations, Nonlinear Analysis, 39 (2000), pp. 135-156.

[10] D. A. French and J. T. King, Analysis of a robust finite element approximation for a parabolic equation with rough boundary data, Math. Comput., 60 (1993), pp. 79-104.

[11] H. Fu and H. Rui, A priori error estimates for optimal control problems governed by transient advection-diffusion equations, J. Sci. Comput., 38 (2009), pp. 290-315.

[12] R. Glowinski, J. L. Lions, Exact and approximate controllability for distributed parameter systems, Acta Numerica, 1996, pp. 159-333.

[13] W. Gong, Error estimates for finite element approximation of parabolic equations involving measure data in space, preprint.

[14] W. Gong and M. Hinze, Error estimates for parabolic optimal control problems with control and state constraints, preprint, Hamburger Beiträge zur Angewandten Mathematik, 2010-13.

[15] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case, Computational Optimization and Application, 30 (2005), pp. 45-63.

[16] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, Optimization with PDE constraints MMTA 23, Springer, 2009.

[17] R. Li, W. Liu, AFEPack available online *http://circus.math.pku.edu.cn/AFEPack*.

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

[19] J. L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications, Springer-Verlag, Berlin, 1972.

[20] W. Liu, N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, Numer. Math., 93 (2003), pp. 497-521.

[21] W. Liu and N. Yan, Adaptive finite element methods for optimal control governed by PDEs. Science press. Beijing, 2008.

[22] M. Luskin and R. Rannacher, On the smoothing property of the Galerkin method for parabolic equations, SIAM J. Numer. Anal., 19 (1982), pp. 93-113.

[23] A. Martínez, C. Rodríguez and M. E. Vázquez-Méndez, Theoretical and numerical analysis of an optimal control problem related to wastewater treatment, SIAM J. Control Optim., 38 (2000), pp. 1534-1553.

[24] D. Meidner, R. Rannacher, and B. Vexler, A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time, Preprint-Nr.: SPP1253-098.

[25] 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., 47 (2008), pp. 1301-1329.

[26] A. M. Ramos, R. Glowinski and J. Periaux, Pointwise control of the Burgers equation and related Nash equilibrium problems: Computational approach, J. Optim. Theory Appl., 112 (2002), pp. 499-516.
[27] R. Rannacher and R. Scott, Some optimal error estimates for piecewise linear finite element approximations, Math. Comput., 38 (1982), pp. 437-445.
[28] R. Scott, Optimal L^∞ estimates for the finite element method on irregular meshes, Math. Comput., 30 (1976), pp. 681-697.
[29] V. Thomée, Galerkin finite element methods for parabolic problems, Springer-Verlag, Berlin, 2006.
[30] J. Zhu and Q. Zeng, A mathematical formulation for optimal control of air pollution, Science in China (D), 46 (10), 994-1002(2003).