Hamburger Beiträge
zur Angewandten Mathematik

Space-time finite element approximation of parabolic optimal control problems

Wei Gong, Michael Hinze, and Zhaojie Zhou

Nr. 2011-22
November 2011

SPACE-TIME FINITE ELEMENT APPROXIMATION OF PARABOLIC OPTIMAL CONTROL PROBLEMS

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

Abstract: In this paper we investigate a space-time finite element approximation of parabolic optimal control problems. The first order optimality conditions are transformed into an elliptic equation of fourth order in space and second order in time involving only the state or the adjoint state in the space-time domain. We derive a priori and a posteriori error estimates for the time discretization of the state and the adjoint state. Furthermore, we also propose a space-time mixed finite element discretization scheme to approximate the space-time elliptic equations, and derive a priori error estimates for the state and the adjoint state. Numerical examples are presented to illustrate our theoretical findings and the performance of our approach.

Key words. Parabolic optimal control problem, space-time finite elements, mixed finite elements, a priori error estimates, a posteriori error estimates.

Subject Classification: 65N15, 65N30.

1. Introduction

Optimal control problems governed by time-dependent partial differential equations play an important role in many practical applications. Their numerical approximation forms a hot research topic and there exist lots of contributions to error analysis and numerical algorithms for time-dependent optimal control problems. For recent works on this topic we refer to, e.g., [16, 20, 21, 22, 23] and the references cited therein. In the present work we build the numerical analysis on reformulations of the optimality conditions as fourth order in space and second order in time elliptic boundary value problems for the state and the adjoint state which are valid under natural regularity assumptions on the data. This approach has recently been used in [25] to tackle parabolic optimal control problems numerically, and was motivated in e.g. [4], where a multi-grid method in the spirit of Hackbusch [7, 8] is proposed to solve parabolic optimal control problems. For a detailed discussion of multigrid methods in the context of optimization problems with PDE constraints we refer the reader to [2]. Multigrid methods are also applied to the numerical solution of optimal control problems with nonlinear PDE systems, see e.g. [15] for their application in flow control.

In this paper we present a discrete formulation of parabolic optimal control problems based on the reformulation of their optimality conditions as second order in time and fourth order in space elliptic boundary value problems. Since time and space have different physical meanings we separate temporal and spatial discretization and put our focus on a priori and a posteriori error analysis for the temporal discretization. In the a posteriori error analysis part we construct residual based error estimators for the time discretization and keep the space variable continuous. The key idea here consists in applying residual-based a posteriori error estimation techniques for two-point boundary value problems to the a posteriori error estimation of space-time elliptic boundary value problems for the state and the adjoint state. Furthermore, we prove a priori error estimates for

* Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146, Hamburg, Germany, and 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.

◇ Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146, Hamburg, Germany, School of Mathematics Sciences, and Shandong Normal University, 250014 Ji'nan, China. Email: zzj534@amss.ac.cn .

temporal semi-discretization with piecewise linear, continuous finite elements, and propose a space-time mixed finite element method to approximate the state and the adjoint state, for which we also prove a priori error estimates. Finally, numerical examples are presented to illustrate the theoretical findings. We note that in [13] an a posteriori space-time finite element approach is presented which also is based on a reformulation of the optimality system as second order in time and fourth order in space elliptic equation.

The outline of this paper is as follows. In section 2 we present the first order optimality conditions for parabolic optimal control problems and derive the space-time domain elliptic systems for the state and the adjoint state. We prove existence and uniqueness of solutions to the space-time elliptic boundary value problems. In section 3 we derive a priori and a posteriori error estimates for the time discretization scheme. Section 4 is devoted to the numerical analysis of the space-time mixed finite element approximations of the state and the adjoint state. Numerical examples are presented in section 5 to illustrate our analytical findings.

Let $\Omega \subset \mathbb{R}^n$ $(1 \leqslant n \leqslant 3)$ be an open bounded domain with sufficiently smooth boundary $\Gamma := \partial\Omega$, $\Omega_T = \Omega \times (0,T]$, $\Sigma_T = \Gamma \times (0,T]$. Throughout this paper we denote by $H^m(\Omega)$ and $H^m(\Omega_T)$ the usual Sobolev space on Ω and Ω_T of integer order $m \geqslant 0$ with norm $\|\cdot\|_{m,\Omega}$ and $\|\cdot\|_{m,\Omega_T}$, respectively. $H^m(\Gamma)$ and $H^m(\Sigma_T)$ are defined accordingly. For $m = 0$ we have $H^0(\Omega) = L^2(\Omega), H^0(\Omega_T) = L^2(\Omega_T)$, $H^0(\Gamma) = L^2(\Gamma)$, and $H^0(\Sigma_T) = L^2(\Sigma_T)$. For the analysis we need $H^{2,1}(\Omega_T) = L^2(0,T;H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ equipped with the norm

$$\|w\|_{2,1,\Omega_T} = \left(\|w\|^2_{L^2(0,T;H^2(\Omega))} + \|w\|^2_{H^1(0,T;L^2(\Omega))}\right)^{\frac{1}{2}}.$$

For the state space we take $X \equiv W(0,T) = \{v \in L^2(0,T;H_0^1(\Omega));\ \frac{\partial v}{\partial t} \in L^2(0,T;H^{-1}(\Omega))\}$. For the control space we take $U := L^2(\Omega_T)$. Throughout the paper c and C denote generic positive constants.

2. Optimal control problem

In this paper we consider the optimal control problem

$$\min_{(y,u)\in X\times U} \quad J(y,u) \quad = \tfrac{1}{2}\|y - y_d\|^2_{0,\Omega_T} + \tfrac{\alpha}{2}\|u\|^2_{0,\Omega_T} \tag{2.1}$$

subject to

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = u & \text{in } \Omega_T, \\ y = 0 & \text{on } \Sigma_T, \\ y(0) = y_0 & \text{in } \Omega, \end{cases} \tag{2.2}$$

where $\alpha > 0$, $y_0 \in H_0^1(\Omega)$, $y_d \in L^2(\Omega_T)$ and $T > 0$ are fixed. The analysis of this optimal control problem is well understood. In e.g. [18] among other things the following theorem is proven.

Theorem 2.1. *The control problem (2.1)-(2.2) admits a unique solution $(y,u) \in X \times U$. The pair (y,u) is the solution of (2.1)-(2.2) if and only if there exists a unique adjoint state $p \in X$ such that the triplet (y,p,u) satisfies the optimality system*

$$\frac{\partial y}{\partial t} - \Delta y = u \quad \text{in } \Omega_T,\ y = 0 \quad \text{on } \Sigma_T,\ y(0) = y_0 \quad \text{in } \Omega, \tag{2.3}$$

$$-\frac{\partial p}{\partial t} - \Delta p = y - y_d \quad \text{in } \Omega_T,\ p = 0 \quad \text{on } \Sigma_T,\ p(T) = 0 \quad \text{in } \Omega, \tag{2.4}$$

$$\alpha u + p = 0, \quad \text{in } \Omega_T. \tag{2.5}$$

For the state y and adjoint state p we have the following regularity results.

Lemma 2.2. *If $y_0 \in H_0^1(\Omega)$ and $y_d \in L^2(\Omega_T)$, we according to* [10] *have*

$$y, p \in H^{2,1}(\Omega_T).$$

Furthermore, if in addition $y_0 \in H^3(\Omega)$ *and the following compatibility conditions*

$$g_0 := y_0 \in H_0^1(\Omega),\ g_1 := u(0) + \Delta g_0 \in H_0^1(\Omega)$$

hold, we by [10] *also have*

$$(a) \qquad y \in L^2(0,T;H^4(\Omega)) \cap H^2(0,T;L^2(\Omega)) \cap H^1(0,T;H^2(\Omega)).$$

Similarly, if we suppose that $y_d \in H^{2,1}(\Omega_T)$ *and that the following compatibility conditions*

$$\widetilde{g}_0 := p(T) \in H_0^1(\Omega),\ \widetilde{g}_1 := y(T) - y_d(T) + \Delta\widetilde{g}_0 \in H_0^1(\Omega)$$

hold, [10] *delivers*

$$(b) \qquad p \in L^2(0,T;H^4(\Omega)) \cap H^2(0,T;L^2(\Omega)) \cap H^1(0,T;H^2(\Omega)).$$

Remark 2.3. *Since we assume that our domain* Ω *is sufficiently smooth, the regularities of the optimal state, the optimal control, and the associated adjoint are limited through the regularities of the initial state* y_0 *and of the desired state* y_d*. By a bootstrap argument we from* (2.3)-(2.5) *infer for the state* y *the regularity*

$$(c) \qquad y \in L^2(0,T;H^6(\Omega)) \cap H^1(0,T;H^4(\Omega)) \cap H^2(0,T;H^2(\Omega)) \cap H^3(0,T;L^2(\Omega)).$$

For the adjoint p *we obtain with the assumption* $y_d \in L^2(0,T;H^4(\Omega)) \cap H^1(0,T;H^2(\Omega)) \cap H^2(0,T;L^2(\Omega))$ *that*

$$(d) \qquad p \in L^2(0,T;H^6(\Omega)) \cap H^1(0,T;H^4(\Omega)) \cap H^2(0,T;H^2(\Omega)) \cap H^3(0,T;L^2(\Omega))$$

holds.

We now show that the optimal state y and the associated adjoint state p under natural regularity assumptions on the data also form solutions to certain $2nd-$order in time and $4th-$order in space elliptic partical differential equations. More precisely, we shall show that y solves

$$(2.6) \qquad \begin{cases} -\dfrac{\partial^2 y}{\partial t^2} + \Delta^2 y + \dfrac{1}{\alpha} y = \dfrac{1}{\alpha} y_d & \text{in } \Omega_T, \\ y = 0 & \text{on } \Sigma_T, \\ \Delta y = 0 & \text{on } \Sigma_T, \\ (\dfrac{\partial y}{\partial t} - \Delta y)(T) = 0 & \text{in } \Omega, \\ y(0) = y_0 & \text{in } \Omega, \end{cases}$$

while p forms a solution to

$$(2.7) \qquad \begin{cases} -\dfrac{\partial^2 p}{\partial t^2} + \Delta^2 p + \dfrac{1}{\alpha} p = -\dfrac{\partial y_d}{\partial t} + \Delta y_d & \text{in } \Omega_T, \\ p = 0 & \text{on } \Sigma_T, \\ \Delta p = y_d & \text{on } \Sigma_T, \\ (\dfrac{\partial p}{\partial t} + \Delta p)(0) = y_d(0) - y_0 & \text{in } \Omega, \\ p(T) = 0 & \text{in } \Omega. \end{cases}$$

We now provide the notions of weak solutions for (2.6) and (2.7), respectively. For this purpose let us define the spaces

$$\begin{aligned} H_0^{2,1}(\Omega_T): &= \Big\{ v \in H^{2,1}(\Omega_T): \ v(0) = 0 \ \text{ in } \ \Omega \Big\}, \\ \widetilde{H}_0^{2,1}(\Omega_T): &= \Big\{ v \in H^{2,1}(\Omega_T): \ v(T) = 0 \ \text{ in } \ \Omega \Big\}, \end{aligned}$$

the bilinear forms

$$\begin{aligned} A_T &: H_0^{2,1}(\Omega_T) \times H_0^{2,1}(\Omega_T) \longrightarrow \mathbb{R}, \\ A_0 &: \widetilde{H}_0^{2,1}(\Omega_T) \times \widetilde{H}_0^{2,1}(\Omega_T) \longrightarrow \mathbb{R}, \end{aligned}$$

as well as linear forms

$$\begin{aligned} L_T &: H_0^{2,1}(\Omega_T) \longrightarrow \mathbb{R}, \\ L_0 &: \widetilde{H}_0^{2,1}(\Omega_T) \longrightarrow \mathbb{R}, \end{aligned}$$

where

$$\begin{aligned} A_T(v_1, v_2) &:= \int_{\Omega_T} (\frac{\partial v_1}{\partial t}\frac{\partial v_2}{\partial t} + \frac{1}{\alpha} v_1 v_2) + \int_{\Omega_T} \Delta v_1 \Delta v_2 + \int_{\Omega} \nabla v_1(T) \nabla v_2(T), \\ A_0(v_1, v_2) &:= \int_{\Omega_T} (\frac{\partial v_1}{\partial t}\frac{\partial v_2}{\partial t} + \frac{1}{\alpha} v_1 v_2) + \int_{\Omega_T} \Delta v_1 \Delta v_2 + \int_{\Omega} \nabla v_1(0) \nabla v_2(0), \\ L_T(v) &:= \int_{\Omega_T} \frac{1}{\alpha} y_d v, \text{ and} \\ L_0(v) &:= \int_0^T \langle -\frac{\partial y_d}{\partial t} + \Delta y_d, v\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} - \int_\Omega (y_d(0) - y_0)v(0) + \int_{\Sigma_T} y_d \nabla v \cdot n. \end{aligned}$$

Definition 2.4. *We call $y \in H^{2,1}(\Omega_T)$ with $y(0) = y_0 \in H_0^1(\Omega)$ a weak solution to* (2.6)*, if y satisfies*

$$A_T(y, v) = L_T(v) \quad \forall\, v \in H_0^{2,1}(\Omega_T). \tag{2.8}$$

Let $y_d \in H^{2,1}(\Omega_T)$. We call $p \in \widetilde{H}_0^{2,1}(\Omega_T)$ a weak solution to (2.7)*, if p satisfies*

$$A_0(p, \phi) = L_0(\phi) \quad \forall\, \phi \in \widetilde{H}_0^{2,1}(\Omega_T). \tag{2.9}$$

In the following we prove existence and uniqueness of solutions to (2.8) and (2.9). For this purpose we equip $H^{2,1}(\Omega_T)$ with the inner product

$$\lfloor v, w \rceil := \int_{\Omega_T} \frac{\partial v}{\partial t}\frac{\partial w}{\partial t} + \int_{\Omega_T} vw + \int_{\Omega_T} \Delta v \Delta w,$$

which induces the norm

$$|||v||| := \Big(\|\frac{\partial v}{\partial t}\|_{0,\Omega_T}^2 + \|\Delta v\|_{0,\Omega_T}^2 + \|v\|_{0,\Omega_T}^2 \Big)^{\frac{1}{2}}.$$

For the two norms $\|\cdot\|_{2,1,\Omega_T}$ and $|||\cdot|||$ we have the following equivalence result (compare [24]).

Lemma 2.5. *The norms $\|\cdot\|_{2,1,\Omega_T}$ and $|||\cdot|||$ are equivalent on $H_0^{2,1}(\Omega_T)$, i.e., there exist positive constants c_1 and c_2 such that*

$$c_1 |||v||| \leqslant \|v\|_{2,1,\Omega_T} \leqslant c_2 |||v|||.$$

Proof. For $v \in H_0^{2,1}(\Omega_T)$ we set $u_v := \frac{\partial v}{\partial t} - \Delta v \in L^2(\Omega_T)$. Then v forms a weak solution to

$$\begin{cases} \dfrac{\partial v}{\partial t} - \Delta v = u_v & \text{in } \Omega_T, \\ v = 0 & \text{on } \Sigma_T, \\ v(0) = 0 & \text{in } \Omega, \end{cases}$$

which depends continuously on u_v, i.e. we have

$$\begin{aligned} \|v\|_{2,1,\Omega_T}^2 &\leqslant C\|u_v\|_{0,\Omega_T}^2 \\ &\leqslant C\|\frac{\partial v}{\partial t} - \Delta v\|_{0,\Omega_T}^2 \\ &\leqslant C(\|\frac{\partial v}{\partial t}\|_{0,\Omega_T}^2 + \|\Delta v\|_{0,\Omega_T}^2) \\ &\leqslant c_2 |||v|||^2. \end{aligned}$$

The estimate

$$|||v|||^2 = \|\frac{\partial v}{\partial t}\|_{0,\Omega_T}^2 + \|\Delta v\|_{0,\Omega_T}^2 + \|v\|_{0,\Omega_T}^2 \leqslant \|v\|_{2,1,\Omega_T}^2$$

follows directly from the definition of $\|v\|_{2,1,\Omega_T}$. This gives the claim. $\square$

Similarly one can prove that the norm $\|\cdot\|_{2,1,\Omega_T}$ is also equivalent to $|||\cdot|||$ on $\widetilde{H}_0^{2,1}(\Omega_T)$. We are now in a position to prove

Theorem 2.6. *There exists a unique weak solution $y \in H^{2,1}(\Omega_T)$ to problem (2.8). If in addition $y_d \in W(0,T)$, (2.9) admits a unique weak solution $p \in \widetilde{H}_0^{2,1}(\Omega_T)$.*

Proof. Since $H^{2,1}(\Omega_T) \hookrightarrow C([0,T]; H^1(\Omega))$,

$$\begin{aligned}\int_\Omega \nabla v_1(T)\nabla v_2(T) &\leqslant \|\nabla v_1(T)\|_{0,\Omega}\|\nabla v_2(T)\|_{0,\Omega}\\ &\leqslant \|v_1\|_{C([0,T];H^1(\Omega))}\|v_2\|_{C([0,T];H^1(\Omega))}\\ &\leqslant C\|v_1\|_{2,1,\Omega_T}\|v_2\|_{2,1,\Omega_T}\end{aligned}$$

holds for $v_1, v_2 \in H^{2,1}(\Omega_T)$. This implies

$$\begin{aligned}&A_T(v_1,v_2)\\ \leqslant\ &\|\frac{\partial v_1}{\partial t}\|_{0,\Omega_T}\|\frac{\partial v_2}{\partial t}\|_{0,\Omega_T} + \frac{1}{\alpha}\|v_1\|_{0,\Omega_T}\|v_2\|_{0,\Omega_T} + \|\Delta v_1\|_{0,\Omega_T}\|\Delta v_2\|_{0,\Omega_T} + C\|v_1\|_{2,1,\Omega_T}\|v_2\|_{2,1,\Omega_T}\\ \leqslant\ &C\|v_1\|_{2,1,\Omega_T}\|v_2\|_{2,1,\Omega_T}\end{aligned}$$

with a positive constant C. Moreover, we have

$$A_T(v,v) \geqslant \int_{\Omega_T}((\frac{\partial v}{\partial t})^2 + \frac{1}{\alpha}v^2) + \int_{\Omega_T}(\Delta v)^2.$$

Lemma 2.5 now yields

$$A_T(v,v) \geqslant C\|v\|^2_{2,1,\Omega_T},$$

which implies that A_T is coercive. Note that L_T is linear and bounded. Therefore, the Lax-Milgram theorem implies that the weak formulation (2.8) admits a unique solution $y \in H^{2,1}(\Omega_T)$.

Similarly we can prove that the bilinear form $A_0(\cdot,\cdot)$ is bounded and coercive , i.e.,

$$A_0(v_1,v_2) \leqslant C\|v_1\|_{2,1,\Omega_T}\|v_2\|_{2,1,\Omega_T}$$

and

$$A_0(v,v) \geqslant C\|v\|^2_{2,1,\Omega_T}.$$

Since L_0 is linear and bounded, i.e.,

$$|L_0(v)| \leqslant C\|v\|_{2,1,\Omega_T},$$

the Lax-Milgram theorem again gives the existence and uniqueness of a solution $p \in \widetilde{H}_0^{2,1}(\Omega_T)$ to the weak formulation (2.9). This completes the proof. $\square$

For our optimal control problem (2.1) we now can prove

Theorem 2.7. *Let $(y,u) \in X \times U$ denote the solution to problem (2.1)-(2.2) with associated adjoint state $p \in X$. Assume that y satisfies* (*a*) *of Lemma 2.2. Then y satisfies* (2.6) *a.e. in space time, and is a weak solution to* (2.6). *If p satisfies* (*b*) *of Lemma 2.2, then p solves* (2.7) *a.e. in space time, and is a weak solution to (2.7).*

Proof. Since the solution y of (2.1)-(2.2) together with adjoint state p satisfy the regularity of Lemma 2.2, we may insert (2.5) into (2.3) and take the derivative with respect to time. This yields

$$\frac{\partial^2 y}{\partial t^2} - \Delta y_t = -\frac{1}{\alpha}p_t.$$

Inserting this equation into the adjoint equation we obtain

$$\frac{\partial^2 y}{\partial t^2} - \Delta y_t = \frac{1}{\alpha}(\Delta p + y - y_d).$$

Now, we use the state equation to replace p in the previous equation. This gives

$$\begin{aligned}\frac{\partial^2 y}{\partial t^2} - \Delta y_t &= \frac{1}{\alpha}(\Delta p + y - y_d)\\ &= \frac{1}{\alpha}(y - y_d) + \Delta(-\frac{\partial y}{\partial t} + \Delta y)\\ &= -\Delta y_t + \Delta^2 y + \frac{1}{\alpha}(y - y_d).\end{aligned}$$

Thus

$$-\frac{\partial^2 y}{\partial t^2} + \Delta^2 y + \frac{1}{\alpha} y = \frac{1}{\alpha} y_d. \tag{2.10}$$

The boundary conditions together with initial value for state variable then read

$$y = 0 \quad \text{on } \Sigma_T,\ y(0) = y_0 \quad \text{in } \Omega. \tag{2.11}$$

From the boundary condition of adjoint state p we obtain

$$0 = p = \alpha(\Delta y - \frac{\partial y}{\partial t}) \quad \text{on } \Sigma_T,$$

which implies

$$\Delta y = 0 \quad \text{on } \Sigma_T. \tag{2.12}$$

Note that $p(T) = 0$ in Ω. Thus

$$0 = p(T) = \alpha(\Delta y - \frac{\partial y}{\partial t})(T) \quad \text{in } \Omega. \tag{2.13}$$

Collecting (2.10)-(2.13) gives (2.6). Therefore, we can conclude that the state y satisfies the space-time elliptic boundary value problem (2.6) a.e. in space-time. Now let $y \in L^2(0,T;H^4(\Omega)) \cap H^2(0,T;L^2(\Omega)) \cap H^1(0,T;H^2(\Omega))$. Then $y \in H^{2,1}(\Omega_T)$ and by Green's formula one can easily prove that y satisfies (2.8), which implies that y is also a weak solution to (2.6).

By similar arguments we can prove that the adjoint state p satisfies the space-time elliptic equations (2.7) a.e. in space-time under the assumption $y_d \in H^{2,1}(\Omega_T)$, and also forms a weak solution to (2.7), where we only have to require $y_d \in W(0,T)$.

□

3. Error estimates for the time discretization scheme

In this section we present a priori and a posteriori error analysis for the temporal discretization, while the space variable is kept continuous.

Let $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$ be a time grid with $k_n = t_n - t_{n-1},\ n = 1, 2, \cdots, M$. Set $I_n = [t_{n-1}, t_n]$ and $k = \max_{1 \leqslant n \leqslant M} k_n$. For the time discretization of the state y we define

$$V_t^k = \{v \in H^{2,1}(\Omega_T);\ v(\cdot)|_{I_n} \in P_1(I_n)\},\quad \overline{V}_t^k = V_t^k \cap H_0^{2,1}(\Omega_T)$$

and consider the semi-discrete problem: Find $y_k \in V_t^k$ with $y_k(0) = y_0$ and

$$A_T(y_k, v_k) = L_T(v_k) \quad \forall\ v_k \in \overline{V}_t^k. \tag{3.1}$$

Since V_t^k is a closed subspace of $H^{2,1}(\Omega_T)$, the Lax-Milgram theorem implies that (3.1) admits a unique solution $y_k \in V_t^k$. For this scheme we have

Theorem 3.1. *Let $y \in H^{2,1}(\Omega_T)$ and $y_k \in V_t^k$ be the solutions to (2.8) and (3.1), respectively. Assume that $y \in H^2(0,T;H^2(\Omega)) \cap W^{2,\infty}(0,T;H^1(\Omega))$. Then we have*

$$\|y - y_k\|_{0,\Omega_T} + k\|y - y_k\|_{H^1(0,T;L^2(\Omega))} \leqslant Ck^2.$$

Proof. From (2.8) and (3.1) we obtain the following error equation:

$$\int_{\Omega_T}\Big(\frac{\partial(y-y_k)}{\partial t}\frac{\partial v_k}{\partial t} \quad + \quad \frac{1}{\alpha}(y-y_k)v_k\Big)+\int_{\Omega_T}\Delta(y-y_k)\Delta v_k$$

$$+\int_{\Omega}\nabla(y-y_k)(T)\nabla v_k(T)=0 \quad \forall\ v_k\in \overline{V}_t^k. \tag{3.2}$$

Let $R_k y\in V_t^k$ denote the temporal Ritz projection ([3]) of y, which is defined by

$$\int_{\Omega_T}\frac{\partial(y-R_ky)}{\partial t}\frac{\partial v_k}{\partial t}=0, \quad \forall\ v_k\in\overline{V}_t^k.$$

Decompose $y-y_k=y-R_ky+R_ky-y_k=\xi_y+\eta_y$. Then we can rewrite (3.2) as

$$\int_{\Omega_T}\Big(\frac{\partial\eta_y}{\partial t}\frac{\partial v_k}{\partial t}+\frac{1}{\alpha}\eta_y v_k\Big)+\int_{\Omega_T}\Delta\eta_y\Delta v_k+\int_{\Omega}\nabla\eta_y(T)\nabla v_k(T)$$

$$= \quad -\int_{\Omega_T}\frac{1}{\alpha}\xi_y v_k-\int_{\Omega}\nabla\xi_y(T)\nabla v_k(T)-\int_{\Omega_T}\Delta\xi_y\Delta v_k. \tag{3.3}$$

Testing (3.3) with $v_k=\eta_y$ leads to

$$\int_{\Omega_T}\Big((\frac{\partial\eta_y}{\partial t})^2+\frac{1}{\alpha}\eta_y^2\Big)+\int_{\Omega_T}\Delta\eta_y^2+\int_{\Omega}\nabla\eta_y(T)^2$$

$$= \quad -\int_{\Omega_T}\frac{1}{\alpha}\xi_y\eta_y-\int_{\Omega}\nabla\xi_y(T)\nabla\eta_y(T)-\int_{\Omega_T}\Delta\xi_y\Delta\eta_y.$$

Using the Young's inequality and error estimates for the Ritz projection ([27]) we obtain

$$\begin{aligned}&\frac{1}{\alpha}\|\eta_y\|_{0,\Omega_T}^2+\|\frac{\partial\eta_y}{\partial t}\|_{0,\Omega_T}^2+\|\Delta\eta_y\|_{0,\Omega_T}^2+\|\nabla\eta_y(T)\|_{0,\Omega}^2\\ \leqslant\ & C(\|\xi_y\|_{0,\Omega_T}^2+\|\nabla\xi_y(T)\|_{0,\Omega}^2+\|\Delta\xi_y\|_{0,\Omega_T}^2)\\ \leqslant\ & Ck^4(\|y\|_{H^2(0,T;L^2(\Omega))}^2+\|y\|_{W^{2,\infty}(0,T;H^1(\Omega))}^2+\|\Delta y\|_{H^2(0,T;L^2(\Omega))}^2).\end{aligned}$$

Then triangle inequality and error estimates for the Ritz projection finally give the estimates

$$\|y-y_k\|_{H^1(0,T;L^2(\Omega))}\leqslant Ck, \quad \|y-y_k\|_{0,\Omega_T}\leqslant Ck^2.$$

□

For the temporal discretization of the adjoint state p we proceed similarly as above. We seek $p_k\in\widetilde{V}_t^k$ such that

$$A_0(p_k,\phi_k)=L_0(\phi_k) \quad \forall\ \phi_k\in\widetilde{V}_t^k \tag{3.4}$$

holds. Here $\widetilde{V}_t^k:=V_t^k\cap\widetilde{H}_0^{2,1}(\Omega_T)$. Then a proof similar to that of the previous theorem gives

Theorem 3.2. *Let $p\in\widetilde{H}_0^{2,1}(\Omega_T)$ and $p_k\in\widetilde{V}_t^k$ be the solutions to (2.9) and (3.4), respectively. Suppose that $p\in H^2(0,T;H^2(\Omega))\cap W^{2,\infty}(0,T;H^1(\Omega))$. Then we have*

$$\|p-p_k\|_{0,\Omega_T}+k\|p-p_k\|_{H^1(0,T;L^2(\Omega))}\leqslant Ck^2.$$

Now we are in the position to derive temporal residual type a posteriori error estimates for y and p. We adopt standard Lagrange interpolation in the a posteriori error estimates since $H^1(0,T)\hookrightarrow C([0,T])$.

Theorem 3.3. *Let $y\in H^{2,1}(\Omega_T)$ and $y_k\in V_t^k$ denote the solutions to (2.8) and (3.1), respectively. Then we obtain*

$$\|y-y_k\|_{2,1,\Omega_T}^2\leqslant C\eta_y^2,$$

where

$$\eta_y^2=\sum_n k_n^2\int_{I_n}\|\frac{1}{\alpha}y_d+\frac{\partial^2y_k}{\partial t^2}-\frac{1}{\alpha}y_k-\Delta^2y_k\|_{0,\Omega}^2+\sum_n\int_{I_n}\|\Delta y_k\|_{0,\Gamma}^2.$$

Proof. Let $e^y = y - y_k$, and let $\pi_k e^y$ denote the standard Lagrange type temporal interpolation of e^y. By (2.8), (3.1), Lemma 2.5 and the definition of Lagrange interpolation we have

$$\begin{aligned}
& c\|y-y_k\|^2_{2,1,\Omega_T} \\
\leqslant\ & \frac{1}{\alpha}\|y-y_k\|^2_{0,\Omega_T} + \|\frac{\partial(y-y_k)}{\partial t}\|^2_{0,\Omega_T} + \|\Delta(y-y_k)\|^2_{0,\Omega_T} \\
\leqslant\ & \int_{\Omega_T}\frac{\partial(y-y_k)}{\partial t}\frac{\partial e^y}{\partial t} + \int_{\Omega_T}\frac{1}{\alpha}(y-y_k)e^y + \int_\Omega \nabla(y-y_k)(T)\nabla e^y(T) + \int_{\Omega_T}\Delta(y-y_k)\Delta e^y \\
=\ & \int_{\Omega_T}\frac{\partial(y-y_k)}{\partial t}\frac{\partial(e^y-\pi_k e^y)}{\partial t} + \int_{\Omega_T}\frac{1}{\alpha}(y-y_k)(e^y-\pi_k e^y) \\
& + \int_{\Omega_T}\Delta(y-y_k)\Delta(e^y-\pi_k e^y) + \int_\Omega \nabla(y-y_k)(T)\nabla(e^y-\pi_k e^y)(T) \\
=\ & \frac{1}{\alpha}\int_{\Omega_T} y_d(e^y-\pi_k e^y) - \int_{\Omega_T}\frac{\partial y_k}{\partial t}\frac{\partial(e^y-\pi_k e^y)}{\partial t} - \int_{\Omega_T}\frac{1}{\alpha}y_k(e^y-\pi_k e^y) \\
& - \int_{\Omega_T}\Delta y_k\Delta(e^y-\pi_k e^y) - \int_\Omega \nabla y_k(T)\nabla(e^y-\pi_k e^y)(T) \\
=\ & \frac{1}{\alpha}\int_{\Omega_T} y_d(e^y-\pi_k e^y) - \int_{\Omega_T}\frac{\partial y_k}{\partial t}\frac{\partial(e^y-\pi_k e^y)}{\partial t} - \int_{\Omega_T}\frac{1}{\alpha}y_k(e^y-\pi_k e^y) - \int_{\Omega_T}\Delta y_k\Delta(e^y-\pi_k e^y)
\end{aligned}$$

Integrating by parts on each time interval yields

$$\begin{aligned}
& c\|y-y_k\|^2_{2,1,\Omega_T} \\
\leqslant\ & \sum_n\int_{I_n}\int_\Omega(\frac{1}{\alpha}y_d + \frac{\partial^2 y_k}{\partial t^2} - \frac{1}{\alpha}y_k - \Delta^2 y_k)(e^y-\pi_k e^y) \\
& + \int_{\Sigma_T}\Delta y_k\nabla(\pi_k e^y - e^y)\cdot n
\end{aligned}$$

Using error estimates for Lagrange interpolation, the trace inequality, and Young's inequality we obtain

$$\begin{aligned}
\|y-y_k\|^2_{2,1,\Omega_T} \quad\leqslant\quad & C\sum_n k_n^2\int_{I_n}\|\frac{1}{\alpha}y_d + \frac{\partial^2 y_k}{\partial t^2} - \frac{1}{\alpha}y_k - \Delta^2 y_k\|^2_{0,\Omega} \\
& + C\sum_n\int_{I_n}\|\Delta y_k\|^2_{0,\Gamma}.
\end{aligned}$$

□

We note that the previous theorem in particular implies

$$\begin{aligned}
\|y-y_k\|^2_{H^1(0,T;L^2(\Omega))} \quad\leqslant\quad & C\sum_n k_n^2\int_{I_n}\|\frac{1}{\alpha}y_d + \frac{\partial^2 y_k}{\partial t^2} - \frac{1}{\alpha}y_k - \Delta^2 y_k\|^2_{0,\Omega} \\
& + C\sum_n\int_{I_n}\|\Delta y_k\|^2_{0,\Gamma}.
\end{aligned}$$

Remark 3.4. *Since $\Delta y = 0$ on Σ_T, the term $\sum_n\int_{I_n}\|\Delta y_k\|^2_{0,\Gamma}$ accounts for the violation of this boundary condition on the time-discrete level.*

For the adjoint state p we proceed similarly.

Theorem 3.5. *Let $p \in \widetilde{H}^{2,1}_0(\Omega_T)$ and $p_k \in \widetilde{V}^k_t$ be the solutions to (2.9) and (3.4), respectively. Then we have*

$$\|p-p_k\|^2_{2,1,\Omega_T} \quad\leqslant\quad C\eta_p^2,$$

where

$$\eta_p^2 = \sum_n k_n^2 \int_{I_n} \| -\frac{\partial y_d}{\partial t} + \Delta y_d + \frac{\partial^2 p_k}{\partial t^2} - \frac{1}{\alpha} p_k - \Delta^2 p_k\|_{0,\Omega}^2 + \sum_n \int_{I_n} \|y_d - \Delta p_k\|_{0,\Gamma}^2.$$

Proof. Let $e^p = p - p_k$, and $\pi_k e^p$ be the standard Lagrange type temporal interpolation of e^p. Using (2.9), (3.4), Lemma 2.5 and the definition of Lagrange interpolation we deduce

$$\begin{aligned}
& c\|p-p_k\|_{2,1,\Omega_T}^2 \\
\leqslant\ & \frac{1}{\alpha}\|p-p_k\|_{0,\Omega_T}^2 + \|\frac{\partial (p-p_k)}{\partial t}\|_{0,\Omega_T}^2 + \|\Delta(p-p_k)\|_{0,\Omega_T}^2 \\
\leqslant\ & \int_{\Omega_T} \frac{\partial (p-p_k)}{\partial t}\frac{\partial e^p}{\partial t} + \int_{\Omega_T}\frac{1}{\alpha}(p-p_k)e^p + \int_\Omega \nabla(p-p_k)(0)\nabla e^p(0) + \int_{\Omega_T}\Delta(p-p_k)\Delta e^p \\
=\ & \int_{\Omega_T} \frac{\partial (p-p_k)}{\partial t}\frac{\partial (e^p-\pi_k e^p)}{\partial t} + \int_{\Omega_T}\frac{1}{\alpha}(p-p_k)(e^p-\pi_k e^p) \\
& + \int_{\Omega_T}\Delta(p-p_k)\Delta(e^p-\pi_k e^p) \\
=\ & \int_{\Omega_T}(-\frac{\partial y_d}{\partial t} + \Delta y_d)(e^p-\pi_k e^p) + \int_{\Sigma_T} y_d \nabla(e^p-\pi_k e^p)\cdot n \\
& - \int_{\Omega_T}\frac{\partial p_k}{\partial t}\frac{\partial (e^p-\pi_k e^p)}{\partial t} - \int_{\Omega_T}\frac{1}{\alpha}p_k(e^p-\pi_k e^p) - \int_{\Omega_T}\Delta p_k \Delta(e^p-\pi_k e^p).
\end{aligned}$$

Integrating by parts on each element yields

$$\begin{aligned}
& c\|p-p_k\|_{2,1,\Omega_T}^2 \\
\leqslant\ & \sum_n \int_{I_n}\int_\Omega (-\frac{\partial y_d}{\partial t} + \Delta y_d + \frac{\partial^2 p_k}{\partial t^2} - \frac{1}{\alpha}p_k - \Delta^2 p_k)(e^p-\pi_k e^p) \\
& + \int_{\Sigma_T}(y_d - \Delta p_k)\nabla(e^p-\pi_k e^p)\cdot n.
\end{aligned}$$

Error estimates of Lagrange interpolation, the trace inequality and Young's inequality then give

$$\begin{aligned}
& \|p-p_k\|_{2,1,\Omega_T}^2 \\
\leqslant\ & C\sum_n k_n^2 \int_{I_n} \| -\frac{\partial y_d}{\partial t} + \Delta y_d + \frac{\partial^2 p_k}{\partial t^2} - \frac{1}{\alpha} p_k - \Delta^2 p_k\|_{0,\Omega}^2 \\
& + C\sum_n \int_{I_n} \|y_d - \Delta p_k\|_{0,\Gamma}^2.
\end{aligned}$$

□

Theorem 3.5 also implies

$$\begin{aligned}
& \|p-p_k\|_{H^1(0,T;L^2(\Omega))}^2 \\
\leqslant\ & C\sum_n k_n^2 \int_{I_n} \| -\frac{\partial y_d}{\partial t} + \Delta y_d + \frac{\partial^2 p_k}{\partial t^2} - \frac{1}{\alpha} p_k - \Delta^2 p_k\|_{0,\Omega}^2 \\
& + C\sum_n \int_{I_n} \|y_d - \Delta p_k\|_{0,\Gamma}^2.
\end{aligned}$$

Remark 3.6. *Similarly as above,* $\sum_n \int_{I_n} \|y_d - \Delta p_k\|_{0,\Gamma}^2$ *accounts for the violation of the boundary condition* $\Delta p = y_d$ *on* Σ_T *on the time-discrete level.*

4. Space-time mixed finite element discretization

Several options can be used to tackle the space-time elliptic problems (2.6) and (2.7) numerically. In order to use piecewise linear, continuous finite elements for the spatial discretization we in this section propose a space-time mixed finite element method to treat (2.6) and (2.7) numerically and prove corresponding a priori error estimates.

In order to derive a mixed formulation of (2.6) let $w := -\Delta y$, where y denotes the unique solution to (2.6). Then (2.6) motivates the following mixed formulation: Find (y, w) which satisfies

$$
\tag{4.1}
\begin{cases}
-\dfrac{\partial^2 y}{\partial t^2} - \Delta w + \dfrac{1}{\alpha} y = \dfrac{1}{\alpha} y_d & \text{in } \Omega_T, \\
\Delta y + w = 0 & \text{in } \Omega_T, \\
y = 0 & \text{on } \Sigma_T, \\
w = 0 & \text{on } \Sigma_T, \\
(\dfrac{\partial y}{\partial t} - \Delta y)(T) = 0 & \text{in } \Omega, \\
y(0) = y_0 & \text{in } \Omega.
\end{cases}
$$

Let

$$
Y := \Big\{ v \in H^1(0,T; H_0^1(\Omega)) : \ v(0) = 0 \text{ in } \Omega \Big\}, \quad W := L^2(0,T; H_0^1(\Omega)).
$$

Then the mixed variational form for (4.1) is to find $y \in H^1(0,T; H_0^1(\Omega))$ satisfying $y(0) = y_0$ and $w \in W$ such that

$$
\tag{4.2}
\begin{cases}
\displaystyle\int_{\Omega_T} (\frac{\partial y}{\partial t}\frac{\partial v}{\partial t} + \frac{1}{\alpha} y v) + \int_{\Omega_T} \nabla w \nabla v + \int_\Omega \nabla y(T) \nabla v(T) = \int_{\Omega_T} \frac{1}{\alpha} y_d v \quad \forall\, v \in Y, \\
\displaystyle -\int_{\Omega_T} \nabla y \nabla \phi + \int_{\Omega_T} w\phi = 0 \quad \forall\, \phi \in W.
\end{cases}
$$

In the following we prove that the pair (y, w) with $w := -\Delta y$ and y the unique smooth solution to (2.6) is a solution to the mixed variational form (4.2), and also that this mixed form has at most one solution $(y, w) \in Y \times W$, so that the unique smooth solution y to (2.6) defines the mixed variational solution.

We start with proving that (4.2) admits at most one solution. Suppose that (y_1, w_1) and (y_2, w_2) are two different solutions to problem (4.2). Then $(\widetilde{y}, \widetilde{w}) = (y_1 - y_2, w_1 - w_2)$ satisfies the following homogeneous system

$$
\tag{4.3}
\begin{cases}
\displaystyle\int_{\Omega_T} (\frac{\partial \widetilde{y}}{\partial t}\frac{\partial v}{\partial t} + \frac{1}{\alpha} \widetilde{y} v) + \int_{\Omega_T} \nabla \widetilde{w} \nabla v + \int_\Omega \nabla \widetilde{y}(T) \nabla v(T) = 0 \quad \forall\, v \in Y, \\
\displaystyle -\int_{\Omega_T} \nabla \widetilde{y} \nabla \phi + \int_{\Omega_T} \widetilde{w}\phi = 0 \quad \forall\, \phi \in W.
\end{cases}
$$

Testing (4.3) with $v = \widetilde{y}$ and $\phi = \widetilde{w}$, and adding the resulting equations yields

$$
\int_{\Omega_T} \Big((\frac{\partial \widetilde{y}}{\partial t})^2 + \frac{1}{\alpha} \widetilde{y}^2 \Big) + \int_\Omega (\nabla \widetilde{y}(T))^2 + \int_{\Omega_T} (\widetilde{w})^2 = 0.
$$

Hence,

$$
\widetilde{y} \equiv 0, \quad \widetilde{w} \equiv 0.
$$

Thus, (4.2) admits at most one solution.

Now let $y \in H^{2,1}(\Omega_T) \cap L^2(0,T; H^3(\Omega))$ be the smooth solution to (2.6). Then $w := -\Delta y \in L^2(0,T; H^1(\Omega))$, and y, w satisfy (4.2), as is shown in the following. For $v \in H^1(0,T; C_0^\infty(\Omega))$ we have from (2.8)

$$
\int_{\Omega_T} (\frac{\partial y}{\partial t}\frac{\partial v}{\partial t} + \frac{1}{\alpha} y v) + \int_{\Omega_T} \nabla w \nabla v + \int_\Omega \nabla y(T) \nabla v(T) = \int_{\Omega_T} \frac{1}{\alpha} y_d v.
$$

According to the density argument we derive the first equation of (4.2). Moreover, we have

$$-\int_{\Omega_T} \nabla y \nabla \phi + \int_{\Omega_T} w\phi = 0 \quad \forall\ \phi \in W$$

holds. Thus (y, w) solves (4.2), and consequently, it is the unique solution to (4.2).

Now we are in a position to consider the discretization of (4.2). We will consider space-time discretization and pure time discretization, respectively.

In the case of space-time discretization, let $\mathscr{T}^h$ be a quasi-uniform partitioning of Ω_T into disjoint regular $(n+1)$-simplices τ or rectangles, so that $\overline{\Omega_T} = \bigcup_{\tau \in \mathscr{T}^h} \bar{\tau}$, where element edges lying on the boundary may be curved. Let h_τ denote the diameter of τ. Set $h = \max_\tau h_\tau$. Associated with $\mathscr{T}^h$ is a finite dimensional subspace V^h of $C(\overline{\Omega_T})$, such that $\chi|_\tau$ are linear or bilinear polynomials for all $\chi \in Y^h$ and $\tau \in \mathscr{T}^h$. Let us set $Y^h = V^h \cap H^1(0, T; H_0^1(\Omega)), Y_0^h = V^h \cap Y$ and $W_0^h = V^h \cap W$ in this case.

In the case of pure time discretization, let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ denote a time grid on $[0, T]$ with grid size $h = \max_n h_n$, where $h_n = t_n - t_{n-1}$. Let us set $\mathscr{T}^h = \{I_n\}_{n=1}^N$ with $I_n = [t_{n-1}, t_n]$ and

$$V^h = \{v \in H^1(0, T; H^1(\Omega);\ v(\cdot)|_{I_n} \in P_1(I_n)\}.$$

We in this case set $Y^h = V^h \cap H^1(0, T; H_0^1(\Omega), Y_0^h = V^h \cap Y$ and $W_0^h = V^h \cap W$.

The discrete approximation to (4.2) is to find $y_h \in Y^h,\ w_h \in W_0^h$ such that $y_h(0) = y_{0h}$ and

$$(4.4)\begin{cases} \displaystyle\int_{\Omega_T} (\frac{\partial y_h}{\partial t}\frac{\partial v_h}{\partial t} + \frac{1}{\alpha} y_h v_h) + \int_{\Omega_T} \nabla w_h \nabla v_h + \int_\Omega \nabla y_h(T) \nabla v_h(T) = \int_{\Omega_T} \frac{1}{\alpha} y_d v_h \ \forall\ v_h \in Y_0^h, \\ \displaystyle -\int_{\Omega_T} \nabla y_h \nabla \phi_h + \int_{\Omega_T} w_h \phi_h = 0 \ \forall\ \phi_h \in W_0^h. \end{cases}$$

holds. Here $y_{0h} =: y_0$ for pure time discretization and $y_{0h} =: \widetilde{y}_0(0)$ for space-time discretization, where $\widetilde{y}_0$ is an approximation of y_0 in Y^h, whose approximation properties can be adjusted to the error bounds which we shall prove below.

Now we are in a position to estimate the errors between the solutions (y, w) of (4.2) and (y_h, w_h) of (4.4).

Theorem 4.1. *Let $(y, w) \in Y \times W$ and $(y_h, w_h) \in Y^h \times W_0^h$ be the solutions to (4.2) and (4.4), respectively. In the case of pure time discretization we for $y \in H^2(0, T; H^4(\Omega)) \cap W^{2,\infty}(0, T; H^1(\Omega))$ have*

$$\|y - y_h\|_{H^1(0,T;L^2(\Omega))} \leqslant Ch, \quad \|y - y_h\|_{0,\Omega_T} \leqslant Ch^2 \quad \text{and} \quad \|w - w_h\|_{0,\Omega_T} \leqslant Ch^2.$$

In the fully discrete case we for $y \in H^2(0, T; H^4(\Omega))$ have

$$\|y - y_h\|_{H^1(0,T;L^2(\Omega))} \leqslant Ch \text{ and } \|w - w_h\|_{0,\Omega_T} \leqslant Ch.$$

Proof. From (4.2) and (4.4) we deduce the following error equations

$$\int_{\Omega_T} \Big(\frac{\partial (y - y_h)}{\partial t}\frac{\partial v_h}{\partial t} + \frac{1}{\alpha}(y - y_h) v_h\Big) + \int_{\Omega_T} \nabla (w - w_h) \nabla v_h$$

$$+ \int_\Omega \nabla (y - y_h)(T) \nabla v_h(T) = 0 \quad \forall\ v_h \in Y_0^h, \tag{4.5}$$

$$-\int_{\Omega_T} \nabla (y - y_h) \nabla \phi_h + \int_{\Omega_T} (w - w_h) \phi_h = 0 \quad \forall\ \phi_h \in W_0^h. \tag{4.6}$$

Let $y^I \in Y^h$, $w^I \in W_0^h$ and set $\psi^h = y^I - y_h$. Then (4.5) yields

$$c\|y^I - y_h\|^2_{H^1(0,T;L^2(\Omega))} + |(y^I - y_h)(T)|^2_{1,\Omega}$$

$$
\begin{aligned}
&\leqslant \int_{\Omega_T} \frac{\partial(y^I-y_h)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y^I-y_h)\psi^h + \int_\Omega \nabla(y^I-y_h)(T)\nabla\psi^h(T) \\
&= \int_{\Omega_T} \frac{\partial(y^I-y)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y^I-y)\psi^h + \int_\Omega \nabla(y^I-y)(T)\nabla\psi^h(T) \\
&\quad + \int_{\Omega_T} \frac{\partial(y-y_h)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y-y_h)\psi^h + \int_\Omega \nabla(y-y_h)(T)\nabla\psi^h(T) \\
&= \int_{\Omega_T} \frac{\partial(y^I-y)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y^I-y)\psi^h + \int_\Omega \nabla(y^I-y)(T)\nabla\psi^h(T) - \int_{\Omega_T}\nabla(w-w_h)\nabla\psi^h \\
&= \int_{\Omega_T} \frac{\partial(y^I-y)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y^I-y)\psi^h + \int_\Omega \nabla(y^I-y)(T)\nabla\psi^h(T) \\
&\quad - \int_{\Omega_T}\nabla(w-w^I)\nabla\psi^h - \int_{\Omega_T}\nabla(w^I-w_h)\nabla\psi^h. \qquad (4.7)
\end{aligned}
$$

Similarly, from equation (4.6) we have

$$
\begin{aligned}
&- \int_{\Omega_T}\nabla(w^I-w_h)\nabla\psi^h \\
&= - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y) - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y-y_h) \\
&= - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y) - \int_{\Omega_T}(w^I-w_h)(w-w_h) \\
&= - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y) - \int_{\Omega_T}(w^I-w_h)(w-w^I) - \int_{\Omega_T}(w^I-w_h)(w^I-w_h) \\
&= - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y) - \int_{\Omega_T}(w^I-w_h)(w-w^I) - \|w^I-w_h\|^2_{0,\Omega_T}. \qquad (4.8)
\end{aligned}
$$

Combining (4.7) and (4.8) we get

$$
\begin{aligned}
&c\|y^I-y_h\|^2_{H^1(0,T;L^2(\Omega))} + \|w^I-w_h\|^2_{0,\Omega_T} + |(y^I-y_h)(T)|^2_{1,\Omega} \\
&= \int_{\Omega_T} \frac{\partial(y^I-y)}{\partial t}\frac{\partial\psi^h}{\partial t} + \frac{1}{\alpha}(y^I-y)\psi^h - \int_{\Omega_T}\nabla(w-w^I)\nabla\psi^h + \int_\Omega \nabla(y^I-y)(T)\nabla\psi^h(T) \\
&\quad - \int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y) - \int_{\Omega_T}(w^I-w_h)(w-w^I) \\
&= \sum_{i=1}^{6} E_i. \qquad (4.9)
\end{aligned}
$$

Now it remains to estimate E_i, $i = 1, \cdots, 6$. We note that these error representations are valid for both, the pure time–discrete as well as for the fully discrete case. We now distinguish these cases and proceed with considering the pure time-discrete case. Let y^I and w^I denote the temporal Ritz projection ([3]) of y and the temporal Lagrange interpolation of w, respectively. Then we have the following error estimates:

$$
|E_1| = |\int_{\Omega_T}(y^I-y)_t(y^I-y_h)_t| = 0, \qquad (4.10)
$$

$$
|E_2| = |\frac{1}{\alpha}\int_{\Omega_T}(y^I-y)(y^I-y_h)| \leqslant \alpha^{-1}Ch^2\|y\|_{H^2(0,T;L^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T}, \qquad (4.11)
$$

$$
|E_6| = |\int_{\Omega_T}(w^I-w_h)(w-w^I)| \leqslant Ch^2\|w\|_{H^2(0,T;L^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T}. \qquad (4.12)
$$

Note that w^I and y^I are continuous w.r.t the space variable. By Green's formula we have

$$\begin{aligned} E_3 &= \int_{\Omega_T} \nabla(w-w^I)\nabla(y^I-y_h) \\ &= -\int_{\Omega_T} \Delta(w-w^I)(y^I-y_h) \end{aligned}$$

and

$$\begin{aligned} E_5 &= \int_{\Omega_T} \nabla(y-y^I)\nabla(w^I-w_h) \\ &= -\int_{\Omega_T} \Delta(y-y^I)(w^I-w_h). \end{aligned}$$

Therefore we deduce

$$|E_3| \leqslant Ch^2\|\Delta w\|_{H^2(0,T;L^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T} \tag{4.13}$$

and

$$|E_5| \leqslant Ch^2\|\Delta y\|_{H^2(0,T;L^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T}. \tag{4.14}$$

It remains to estimate E_4. Following [27] we obtain

$$\begin{aligned} |E_4| &= |\int_{\Omega} \nabla(y^I-y)(T)\nabla\psi^h(T)| \\ &\leqslant Ch^2\|y\|_{W^{2,\infty}(0,T;H^1(\Omega))}|\psi^h(T)|_{1,\Omega}. \end{aligned} \tag{4.15}$$

Combining (4.10)-(4.15) we deduce

$$\begin{aligned} &\|y^I-y_h\|^2_{H^1(0,T;L^2(\Omega))} + \|w^I-w_h\|^2_{0,\Omega_T} + |(y^I-y_h)(T)|^2_{1,\Omega} \\ \leqslant\ & C\alpha^{-1}h^2\|y\|_{H^2(0,T;L^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T} + Ch^2\|\Delta w\|_{H^2(0,T;L^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T} \\ &+Ch^2\|y\|_{W^{2,\infty}(0,T;H^1(\Omega))}|(y^I-y_h)(T)|_{1,\Omega} + Ch^2\|\Delta y\|_{H^2(0,T;L^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T} \\ &+Ch^2\|w\|_{H^2(0,T;L^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T}. \end{aligned}$$

Applications of the triangle inequality and of Young's inequality, combined with error estimates for the Ritz projection and Lagrange interpolation, lead to the following estimate

$$\|y-y_h\|_{H^1(0,T;L^2(\Omega))} \leqslant Ch, \quad \|y-y_h\|_{0,\Omega_T} \leqslant Ch^2$$

and

$$\|w-w_h\|_{0,\Omega_T} \leqslant Ch^2.$$

Now we are in a position to derive the error estimates for the space-time discretization case. Let $y^I = R_h y \in Y^h$ and $w^I = R_h w \in W_0^h$ denote the Ritz projections ([3]), which are defined as follows:

$$\int_{\Omega_T} \nabla(\varphi - R_h\varphi)\nabla v_h = 0, \quad \forall\ v_h \in W_0^h$$

for either $\varphi = y$ or $\varphi = w$. Then we have

$$|E_1| = |\int_{\Omega_T} \frac{\partial(y^I-y)}{\partial t}\frac{\partial(y^I-y_h)}{\partial t}| \leqslant Ch\|y\|_{H^2(0,T;H^1(\Omega))}\|\frac{\partial(y^I-y_h)}{\partial t}\|_{0,\Omega_T}, \tag{4.16}$$

$$|E_2| = |\frac{1}{\alpha}\int_{\Omega_T}(y^I-y)(y^I-y_h)| \leqslant \alpha^{-1}Ch^2\|y\|_{H^2(0,T;H^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T}, \tag{4.17}$$

$$|E_3| = |\int_{\Omega_T} \nabla(w-w^I)\nabla(y^I-y_h)| = 0, \tag{4.18}$$

$$|E_4| = |\int_{\Omega} \nabla(y^I-y)(T)\nabla(y^I-y_h)(T)| \leqslant Ch\|y\|_{H^2(0,T;H^2(\Omega))}|(y^I-y_h)(T)|_{1,\Omega}, \tag{4.19}$$

$$(4.20)\quad |E_5| \;=\; |\int_{\Omega_T}\nabla(w^I-w_h)\nabla(y^I-y)| = 0,$$

$$(4.21)\quad |E_6| \;=\; |\int_{\Omega_T}(w^I-w_h)(w-w^I)| \leqslant Ch^2\|w\|_{H^2(0,T;H^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T}.$$

From (4.16)-(4.21) we deduce

$$\begin{aligned}
&\|y^I-y_h\|^2_{H^1(0,T;L^2(\Omega))}+\|w^I-w_h\|^2_{0,\Omega_T}+|(y^I-y_h)(T)|^2_{1,\Omega}\\
\leqslant\; & Ch\|y\|_{H^2(0,T;H^2(\Omega))}\|\frac{\partial(y^I-y_h)}{\partial t}\|_{0,\Omega_T}+C\alpha^{-1}h^2\|y\|_{H^2(0,T;H^2(\Omega))}\|y^I-y_h\|_{0,\Omega_T}\\
&+Ch\|y\|_{H^2(0,T;H^2(\Omega))}|(y^I-y_h)(T)|_{1,\Omega}+Ch^2\|w\|_{H^2(0,T;H^2(\Omega))}\|w^I-w_h\|_{0,\Omega_T}.
\end{aligned}$$

Several applications of the triangle inequality and of Young's inequality combined with error estimates for the Ritz projection lead to

$$\|y-y_h\|_{H^1(0,T;L^2(\Omega))}\leqslant Ch \quad\text{and}\quad \|w-w_h\|_{0,\Omega_T}\leqslant Ch.$$

□

Remark 4.2. *(Superconvergence, see* [28] *for results and details) If we could expect superconvergence properties on uniform structured meshes (for example, on rectangular meshes, on three directional triangular meshes in two spatial dimensions, and on uniform brick meshes in three spatial dimensions), in the fully discrete case we expect the improved estimates*

$$\|y-y_h\|_{0,\Omega_T}\leqslant Ch^2|\log h| \quad\textit{and}\quad \|w-w_h\|_{0,\Omega_T}\leqslant Ch^2|\log h|.$$

In the following we consider the space-time mixed finite element approximation of the adjoint state p. Let $\vartheta := -\Delta p$, where p denotes the unique solution to (2.7). Then (2.7) motivates the following mixed formulation for the adjoint state p:

$$(4.22)\quad \begin{cases}
-\dfrac{\partial^2 p}{\partial t^2}-\Delta\vartheta+\dfrac{1}{\alpha}p=-\dfrac{\partial y_d}{\partial t}+\Delta y_d & \text{in } \Omega_T,\\
\vartheta+\Delta p=0 & \text{in } \Omega_T,\\
p=0 & \text{on } \Sigma_T,\\
-\vartheta=y_d & \text{on } \Sigma_T,\\
(\dfrac{\partial p}{\partial t}+\Delta p)(0)=y_d(0)-y_0 & \text{in } \Omega,\\
p(T)=0 & \text{in } \Omega.
\end{cases}$$

Let

$$\begin{aligned}
\widetilde{Y}:\; &=\; \Big\{p\in H^1(0,T;H^1_0(\Omega)):\; p(T)=0 \text{ in } \Omega\Big\}.\\
\widetilde{W}:\; &=\; L^2(0,T;H^1(\Omega)).
\end{aligned}$$

The mixed variational formulation related to (4.22) is to find $p\in\widetilde{Y},\ \vartheta+y_d\in W$ such that:

$$(4.23)\quad \begin{cases}
\displaystyle\int_{\Omega_T}(\frac{\partial p}{\partial t}\frac{\partial v}{\partial t}+\frac{1}{\alpha}pv)+\int_{\Omega_T}\nabla\vartheta\nabla v+\int_{\Omega}\nabla p(0)\nabla v(0)\\
\displaystyle=\int_{\Omega_T}(-\frac{\partial y_d}{\partial t}+\Delta y_d)v-\int_{\Omega}(y_d(0)-y_0)v(0)\quad \forall\ v\in\widetilde{Y},\\
\displaystyle-\int_{\Omega_T}\nabla p\nabla\phi+\int_{\Omega_T}\vartheta\phi=0\quad \forall\ \phi\in W.
\end{cases}$$

According to (4.23) the mixed finite element approximation of the adjoint state p is to find $p_h \in \widetilde{Y}_0^h$, $\vartheta_h \in V^h$ and $\vartheta_h|_\Gamma = -\widetilde{y}_d$ such that

$$
(4.24) \qquad \left\{
\begin{aligned}
&\int_{\Omega_T} (\frac{\partial p_h}{\partial t}\frac{\partial v_h}{\partial t} + \frac{1}{\alpha} p_h v_h) + \int_{\Omega_T} \nabla\vartheta_h \nabla v_h + \int_\Omega \nabla p_h(0)\nabla v_h(0) \\
&= \int_{\Omega_T} (-\frac{\partial y_d}{\partial t} + \Delta y_d) v_h - \int_\Omega (y_d(0) - y_0) v_h(0), \quad \forall v_h \in \widetilde{Y}_0^h, \\
&\qquad\qquad - \int_{\Omega_T} \nabla p_h \nabla \phi_h + \int_{\Omega_T} \vartheta_h \phi_h = 0, \qquad \forall \phi_h \in W_0^h,
\end{aligned}
\right.
$$

holds, where $\widetilde{Y}_0^h = V^h \cap \widetilde{Y}$ and $\widetilde{y}_d \in V^h$ is an approximation of y_d.

By arguments similar to those applied in the analysis of the state y we get

Theorem 4.3. *Let $(p, \vartheta) \in \widetilde{Y} \times \widetilde{W}$ and $(p_h, \vartheta_h) \in \widetilde{Y}_0^h \times V^h$ denote the solutions to (4.23) and (4.24), respectively. In the pure time–discrete case, we for $p \in H^2(0,T;H^4(\Omega)) \cap W^{2,\infty}(0,T;H^1(\Omega))$ have*

$$\|p - p_h\|_{H^1(0,T;L^2(\Omega))} \leqslant Ch, \quad \|p - p_h\|_{0,\Omega_T} \leqslant Ch^2 \quad \text{and} \quad \|\vartheta - \vartheta_h\|_{0,\Omega_T} \leqslant Ch^2.$$

In the fully discrete case we for $p \in H^2(0,T;H^4(\Omega))$ have

$$\|p - p_h\|_{H^1(0,T;L^2(\Omega))} \leqslant Ch \quad \text{and} \quad \|\vartheta - \vartheta_h\|_{0,\Omega_T} \leqslant Ch.$$

Remark 4.4. *In allignement to Remark 4.2 we for the fully discrete space in the case of superconvergence–meshes expect the estimates*

$$\|p - p_h\|_{0,\Omega_T} \leqslant Ch^2|\log h| \quad \text{and} \quad \|\vartheta - \vartheta_h\|_{0,\Omega_T} \leqslant Ch^2|\log h|.$$

Remark 4.5. *By exploiting the respective second identites in (4.2) and (4.4), the estimates of Theorem 4.1 also yield*

$$\|y - y_h\|_{H^1(\Omega_T)} \leqslant Ch$$

in the pure time–discrete case, and

$$\|y - y_h\|_{H^1(\Omega_T)} \leqslant Ch^{\frac{1}{2}}$$

for space-time discretization. Furthermore, we can expect

$$\|y - y_h\|_{H^1(\Omega_T)} \leqslant Ch\sqrt{|\log h|}$$

for space-time discretization on uniform meshes by using the results of Remark 4.4. We note that similar results also apply to the error estimates for the adjoint state p.

5. Numerical examples

In this section, we will carry out some numerical examples to confirm our theoretical results. Firstly, we test the convergence order of the space-time mixed finite element scheme presented in Section 4. Secondly, we investigate the effectiveness of the temporal a posteriori error estimates η_y and η_p developed in Section 3. In both cases a space-time mixed finite element method is used to solve the space-time elliptic equations. The adaptive procedure is built upon an initial time mesh, together with a refinement strategy based on the local a posteriori error estimator η_{ϕ,I_n} ($\phi = y$ or p) related to the interval I_n together with bulk marking, i.e. all intervalls I_n satisfying

$$\eta_{\phi,I_n} \geqslant \theta \max\{\eta_{\phi,I_n}\},$$

are refined, where $\theta \in (0,1)$. The effectivity index EI is defined as

$$EI = \frac{\eta_\phi}{\|\phi - \phi_k\|_{H^1(0,T;L^2(\Omega))}},$$

where $\phi = y$ or p.

TABLE 1. Errors in y and w for Example 5.1 on a uniform rectangular mesh with fixed space mesh size.

Dof	$\|w - w_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{H^1(0,T;L^2(\Omega))}$	order
2565	0.130512029211	\	0.013224854142	\	0.362590027287	\
4617	0.027382620124	2.2529	0.002775737620	2.2523	0.179149675402	1.0172
8721	0.006451392985	2.0856	0.000654996906	2.0833	0.089179789894	1.0064
16929	0.001594433608	2.0166	0.000162900612	2.0075	0.044535562504	1.0018

TABLE 2. Errors of state y and w for Example 5.1 on a uniform rectangle mesh.

Dof	$\|w - w_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{H^1(0,T;L^2(\Omega))}$	order
25	0.238351935729	\	0.047527619881	\	0.365097724548	\
81	0.058703812693	2.0216	0.012226170429	1.9588	0.179271813135	1.0261
289	0.014624795001	2.0050	0.003078755763	1.9896	0.089187811733	1.0072
1089	0.003657974813	1.9993	0.000771587704	1.9964	0.044536635124	1.0019

TABLE 3. Errors of state y and w for Example 5.1 on general triangle meshes.

Dof	$\|w - w_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{H^1(0,T;L^2(\Omega))}$	order
141	0.038492500552	\	0.008262001121	\	0.200428019476	\
521	0.015827469735	1.3599	0.002168764945	2.0467	0.099641508687	1.0697
2001	0.007415718014	1.1268	0.000556931706	2.0205	0.049566630052	1.0378
7841	0.003527879682	1.0879	0.000141382932	2.0077	0.024676235560	1.0214

TABLE 4. Errors of state y and w for Example 5.1 on uniform triangle meshes.

Dof	$\|w - w_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{0,\Omega_T}$	order	$\|y - y_h\|_{H^1(0,T;L^2(\Omega))}$	order
25	0.586731483264	\	0.097380683188	\	0.623236266193	\
81	0.187433063539	1.6463	0.029508172382	1.7225	0.315028169663	0.9843
289	0.060557215241	1.6300	0.007845736087	1.9111	0.155997022609	1.0140
1089	0.021473108657	1.4958	0.001991686518	1.9779	0.077531227695	1.0087

Example 5.1. *Let $\Omega = [0,1]$, $T = 1$ so that $\Omega_T = (0,1) \times (0,1)$. We choose an exact solution (y,u) of our optimal control problem (2.1) together with an exact adjoint state p as follows*

$$y(x,t) = sin(\pi x)cos(\pi t), \quad p(x,t) = sin(\pi x)sin(\pi t), \quad u(x,t) = -sin(\pi x)sin(\pi t).$$

Here, the corresponding desired state y_d and the right hand side f have to be chosen accordingly.

We first consider the behavior of the error with respect to time. The domain is triangulated into rectangles and bilinear finite elements are applied to approximate the state y. The uniform space mesh has fixed size 1/512, while the time mesh is refined uniformly. In Table 1 the results for the experimental order of convergence are presented for the errors $\|y-y_h\|_{H^1(0,T;L^2(\Omega))}$, $\|w-w_h\|_{0,\Omega_T}$ and $\|y-y_h\|_{0,\Omega_T}$. We find errors of order $O(h)$, $O(h^2)$ and $O(h^2)$, respectively, which is in agreement with the theoretical analysis of Theorem 4.1.

Next, in Table 2 we present the error of the states under uniform refinement with respect to a uniform rectangular space-time mesh. In this case, superconvergence properties may be expected, compare Remark 4.2. In fact, we observe $O(h^2)$ convergence for $\|w - w_h\|_{0,\Omega_T}$ and also for $\|y - y_h\|_{0,\Omega_T}$. In the subsequent tables 3 and 4 we present numerical results for the space–time discretization on general triangle meshes, and on uniform triangle meshes, respectively. One can see that we for general triangle meshes obtain the convergence predicted by Theorem 4.1, whereas on uniform triangle meshes we obtain an error of order $O(h^{3/2})$.

TABLE 5. Errors of state y for Example 5.2 on uniform and adaptive meshes with the same fixed space mesh size and $\varepsilon = 1.0e-003$.

Uniform	$\|y-y_h\|_{0,\Omega_T}$	$\|w-w_h\|_{0,\Omega_T}$	Adaptive	$\|y-y_h\|_{0,\Omega_T}$	$\|w-w_h\|_{0,\Omega_T}$
2121	0.1847	1.8231	2121	0.0158	0.1557
4141	0.1283	1.2671	4343	0.0011	0.0108
8181	0.0755	0.7451	8181	7.2029e-004	0.0068
10201	0.0555	0.5481	10605	2.7602e-004	0.0022

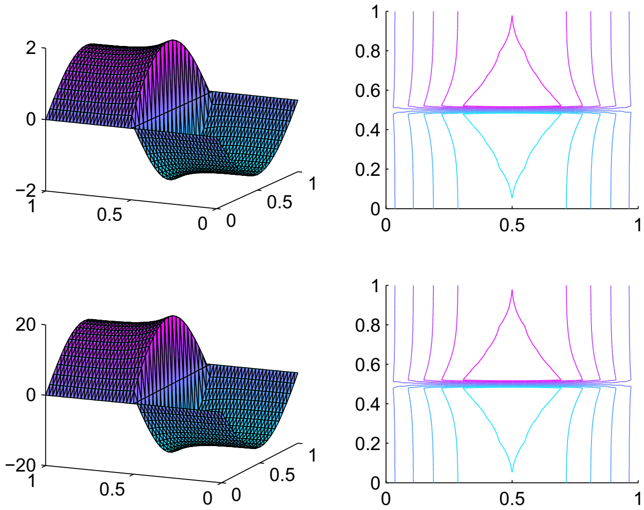

FIGURE 1. The surfaces and contour lines for the states y_h and w_h on uniform meshes for Example 5.2 with $\varepsilon = 1.0e-3$.

Example 5.2. *In this example we investigate the effectiveness of the a posteriori error estimator for the state y presented in Theorem 3.3. Let $\Omega_T = (0,1)\times(0,1)$. The exact solutions are chosen as*

$$y(x,t) = sin(\pi x)atan((t-1/2)/\varepsilon), \quad p(x,t) = sin(\pi x)sin(\pi t), \quad u(x,t) = -sin(\pi x)sin(\pi t)$$

where the corresponding desired state y_d and right hand side f have to be chosen accordingly.

The state y developes an interior layer at $t = 1/2$ for small ε. We use the a posteriori error estimator η_y of Theorem 3.3 to construct the adaptive mesh for the state y.

The surfaces and contour lines for the state y on the uniform and on the adaptive mesh are presented in Figure 1 and 2, respectively. The errors of the state y on uniform and adaptive meshes are displayed in Table 5. The error reduction for the state y on the uniform and on the adaptive mesh is shown in Figure 4. We observe that the adaptive mesh largely improves the quality of the numerical solutions. Figure 3 displays the adaptive mesh and documents the behaviors of the error estimators. We deduce, that our a posteriori error indicator η_y captures the singular layer exactly, with an effectivity index EI close to 2.

All the numerical results clearly indicate that the a posteriori error estimator for the state y is reliable and efficient, and that the use of adaptive meshes can heavily improve the behavior of numerical solutions and also save substantially calculation time.

Example 5.3. *In this example we investigate the behavior of the a posteriori error estimator for the adjoint state p developed in Theorem 3.5. Again $\Omega_T = (0,1)\times(0,1)$. The exact solutions are*

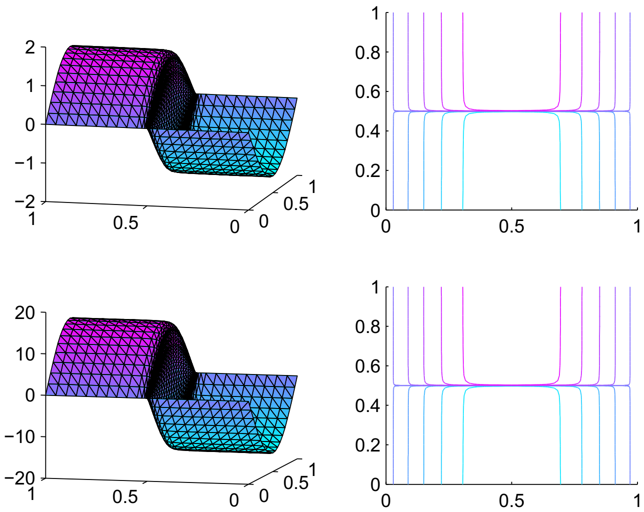

FIGURE 2. The surfaces and contour lines of the states y_h and w_h on adaptive meshes for Example 5.2 with $\varepsilon = 1.0e-3$.

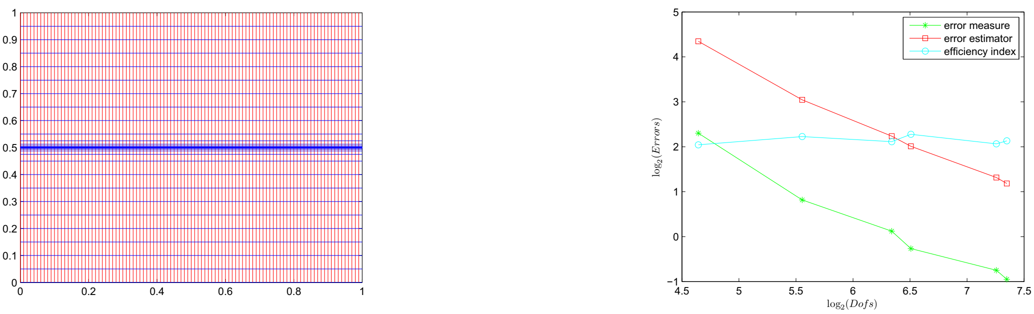

FIGURE 3. The adaptive mesh (left) and the performance of the a posteriori error estimator (right) for the state y.

TABLE 6. Values of η_p^i and η_p^b for Example 5.3 on time–adaptive meshes with a fixed space mesh (left with size $\frac{1}{100}$, right with size $\frac{1}{150}$) and $\varepsilon = 1.0e-004$.

Adaptive	η_p^i	η_p^b	Adaptive	η_p^i	η_p^b
3200	2.4717	3.4169e-004	6600	1.5436	1.2586e-004
5000	1.0833	4.8034e-005	8607	0.9511	4.3404e-005
7100	0.6553	2.7031e-005	10268	0.7285	2.9835e-005
9600	0.4549	2.2352e-005	14647	0.4605	2.2354e-005

chosen as

$$y(x,t) = sin(\pi x)sin(\pi t), \quad p(x,t) = x(x-1)(t - \frac{e^{(t-1)/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}}), \quad u(x,t) = -p(x,t)$$

where the desired state y_d and right hand side f have to be chosen accordingly.

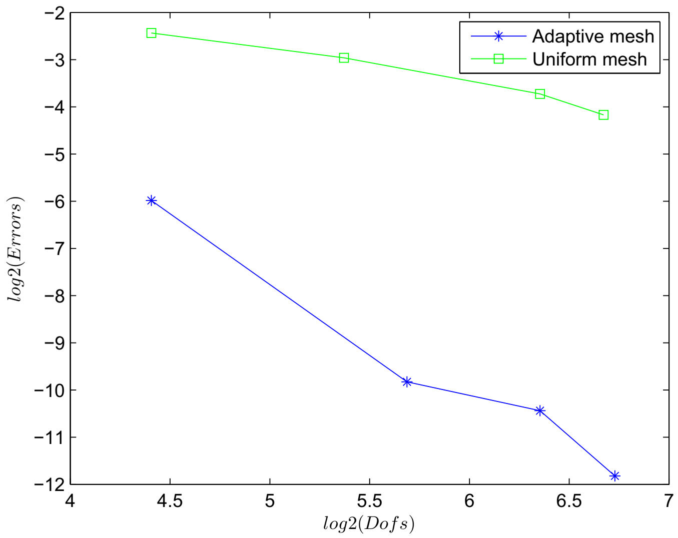

FIGURE 4. The error reduction for the state y on uniform and adaptive meshes.

TABLE 7. Errors of adjoint state p for Example 5.3 on uniform and adaptive time–meshes with a fixed space mesh and $\varepsilon = 1.0e-003$.

Uniform	$\|p-p_h\|_{0,\Omega_T}$	$\|\vartheta-\vartheta_h\|_{0,\Omega_T}$	Adaptive	$\|p-p_h\|_{0,\Omega_T}$	$\|\vartheta-\vartheta_h\|_{0,\Omega_T}$
2626	0.0200	0.2196	2727	4.6762e-005	4.8235e-004
4141	0.0175	0.1873	4343	1.7571e-005	1.4911e-004
6161	0.0124	0.1337	6161	1.4466e-005	1.1573e-004
9191	0.0087	0.0955	9090	1.0441e-005	6.7456e-005

The adjoint state p in this example developes a boundary layer at $t = 1$ if ε is small. We use the a posteriori error estimator η_p of Theorem 3.5 to construct adaptive time meshes for the adjoint state p. We denote by η_p^i and η_p^b, respectively the first and second contribution of the a posteriori error estimator η_p. The values of η_p^i and η_p^b on adaptive times meshes on a fixed space mesh are presented in Table 6. We observe that η_p^i dominates the a posteriori error estimator η_p and that the contribution of η_p^b is, as expected, negligible.

The surfaces and contour lines for the adjoint state p on uniform and adaptive meshes are shown in Figure 5 and 6. The errors of the adjoint p on uniform and adaptive meshes are presented in Table 7. The error reduction for the adjoint state p on uniform and adaptive meshes is documented in Figure 8. We can conclude that the use of adaptive meshes improves the numerical solutions and saves substantially computation time.

The adaptive mesh and the behaviors of errors and error estimators are displayed in Figure 7. We see that the a posteriori error estimator η_p exactly captures the singular boundary layer with effectivity index EI of approximately 2.

ACKNOWLEDGMENTS

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

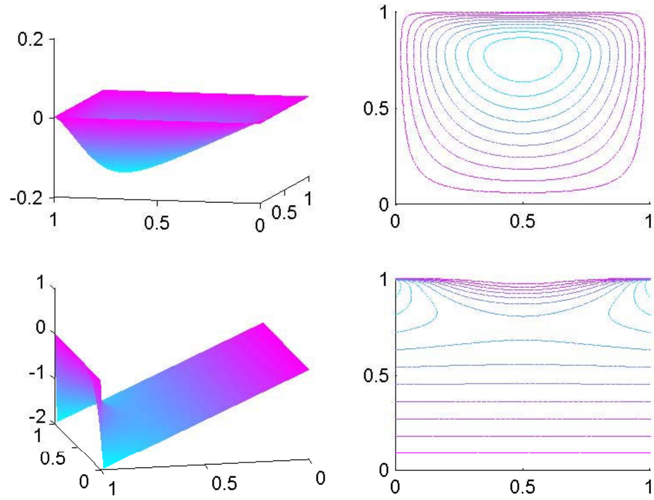

FIGURE 5. The surfaces and contour lines of the adjoint state p_h (top) and of ϑ_h (bottom) on uniform a uniform mesh for Example 5.3 with $\varepsilon = 1.0e-4$.

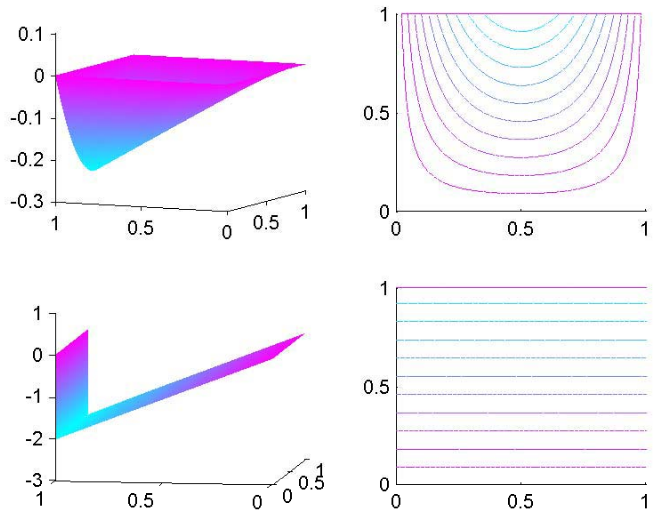

FIGURE 6. The surfaces and contour lines of p_h (top) and ϑ_h (bottom) on adaptive meshes for Example 5.3 with $\varepsilon = 1.0e-4$.

of Mathematics, University of Hamburg for the hospitality and support. Finally, the second author acknowledges support of the Mathematisches Forschungsinstitut Oberwolfach within the Research in Pairs program in 2010. The a posteriori error control concept proposed in the present paper was one of the results of many fruitful discussions with Michael Hintermller from the Humboldt Universitt, and Ronald H.W. Hoppe from the Universitt Augsburg during this stay.

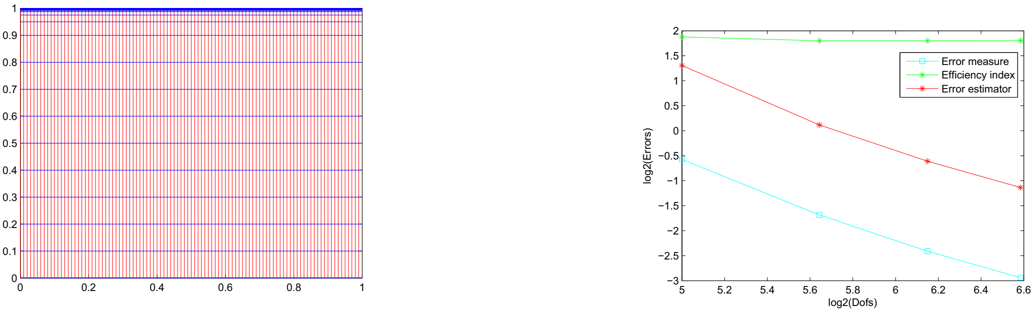

FIGURE 7. The adaptive mesh (left) and the performance of the a posteriori error estimator (right) for the adjoint state p.

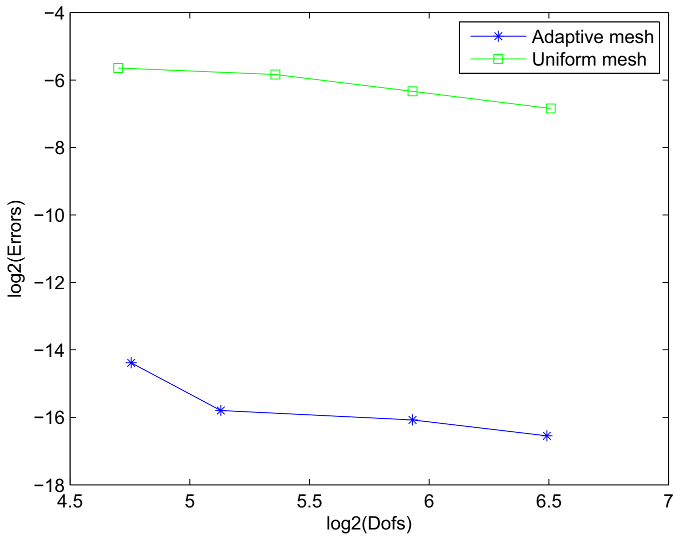

FIGURE 8. The error reduction for adjoint state p on uniform and adaptive meshes.

REFERENCES

[1] S. Balasundaram and P. K. Bhattacharyya, A mixed finite element method for fourth order elliptic equations with variable coefficients, Comput. Math. Appl., 10 (3), 245–256 (1984).

[2] A. Borzì and V. Schulz, Multigrid methods for PDE optimization, SIAM Review, 51, 361-395 (2009).

[3] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics, Vol. 15, Springer-Verlag, New York (2008).

[4] G. Büttner, Ein Mehrgitterverfahren zur optimalen Steuerung parabolischer Probleme, PhD thesis, Fakultät II-Mathematik und Naturwissenschaften der Technischen Universität Berlin, 2004.

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

[6] Ph. Clément, Approximation by finite element functions using local regularization, RAIRO Anal. Numer., Vol. 9, 77-84 (1975).

[7] Hackbusch, W.: A Numerical Method for Solving Parabolic Equations With Opposite Orientations. Computing 20, pp. 229-240, 1978.

[8] Hackbusch, W.: On the fast solving of parabolic boundary control problems. SIAM J.Control. Optim. 17 (1979), pp. 231-244, 1979.

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

[10] L. C. Evans, Partial differential equations, Second edition, Graduate Studies in Mathematics, 19, American Mathematical Society, Providence, RI, 2010.
[11] T. Geveci, On the approximation of the solution of an optimal control problem governed by an elliptic equation, RAIRO Anal. Numer., Vol. 13, 313-328 (1979).
[12] W. Gong and M. Hinze, Error estimates for parabolic optimal control problems with control and state constraints, Hamburger Beiträge zur Angewandten Mathematik 2010-13 (2010).
[13] M.Hintermller, M. Hinze, R. H.W. Hoppe, F. A.M. Ibrahim, and Y. Iliash, Adaptive space-time finite element approximations of parabolic optimal control problems. In preparation.
[14] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case, Comput. Optim. Appl., Vol. 30, 45-63 (2005).
[15] M. Hinze, M. Köster and S. Turek, A space-time multigrid solver for distributed control of the time-dependent Navier-Stokes system, Priority Programme 1253, Preprint-Nr.: SPP1253-16-02 (2008).
[16] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, Optimization with PDE constraints, MMTA 23, Springer, 2009.
[17] A. Kröner, K. Kunisch and B. Vexler, Semismooth Newton methods for optimal control of the wave equation with control constraints, SIAM J. Control Optim., Vol. 49 (2), 830-858 (2011).
[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, H. Ma, T. Tang, N. Yan, A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations, SIAM J. Numer. Anal., Vol. 42, 1032-1061 (2004).
[21] W. Liu and N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, Numer. Math., Vol. 93, 497-521 (2003).
[22] D. Meidner and B. Vexler, Adaptive space-time finite element methods for parabolic optimization problems, SIAM J. Control Optim., Vol. 46 (1), 116-142 (2007).
[23] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints, SIAM J. Control Optim., Vol. 47, 1150-1177 (2008).
[24] I. Neitzel, U. Prüfert and T. Slawig, A smooth regularization of the projection formula for constrained parabolic optimal control problems, Numerical Functional Analysis and Optimization, to appear, 2011.
[25] I. Neitzel and F. Tröltzsch, On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints, ESAIM: Control Optim. Calc. Var., Vol. 15, 426-453 (2009).
[26] L. R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp., Vol. 54, 483-493 (1990).
[27] L. Wahlbin, On maximum norm error estimates for Galerkin approximations to one-dimensional second order parabolic boundary value problems, SIAM J. Numer. Anal., Vol. 12, 177-182 (1975).
[28] N. Yan, Superconvergence analysis and a posteriori error estimation in finite element methods, Science Press, Beijing (2008).