Hamburger Beiträge
zur Angewandten Mathematik

Variational discretization for optimal control governed by convection dominated diffusion equations

Michael Hinze, Ningning Yan and Zhaojie Zhou

Nr. 2008-04
March 2008

Variational discretization for optimal control governed by convection dominated diffusion equations

Michael Hinze
Department of Mathematics, University of Hamburg, Germany
Ningning Yan and Zhaojie Zhou
Institute of Systems Science, Academy of Mathematics and Systems Science
Chinese Academy of Science, China

Abstract

In this paper, we study variational discretization for the constrained optimal control problem governed by convection dominated diffusion equations, where the state equation is approximated by the edge stabilization Galerkin method. A priori error estimates are derived for the state, the adjoint state and the control. Moreover, residual type a posteriori error estimates in the L^2-norm are obtained. Finally, two numerical experiments are presented to illustrate the theoretical results.

Key words: constrained optimal control problem, convection dominated diffusion equation, edge stabilization Galerkin method, variational discretization, a priori error estimate, a posteriori error estimate.
Subject Classification: 65N30.

1 Introduction

Optimal control problem governed by convection dominated diffusion equations arises in many science and engineering applications. Recently, extensive research has been carried out on various theoretical aspects of optimal control problems governed by convection diffusion and convection dominated equations, see, for example, [2], [3], [10], [29].

It is well known that the standard finite element discretizations applied to convection dominated diffusion problems lead to strongly oscillations when layers are not properly resolved. To stabilize this phenomenon, several well-established techniques have been proposed and analyzed, for example, the streamline diffusion finite element method [16], residual free bubbles [4], and the discontinuous Galerkin method [18]. Drawing on earlier ideas by Douglas and Dupont [11], Burman and Hansbo proposed an edge stabilization Galerkin method to approximate the convection dominated diffusion equations in paper [5]. The method uses least square stabilization of the gradient jumps across element edges, and can be seen as a continuous, higher order interior penalty method. The analysis of

[0]The research of the first author was supported by the Priority Programme 1253 entitled Optimization with pde constraints, and is sponsored by the German Research Foundation under Grant HI 689/5-1. The research of the second and third author was supported by National Natural Science Foundation of China (No. 60474027 and 10771211) and the National Basic Research Program under the Grant 2005CB32170X.

edge stabilization Galerkin methods has been extended to the Stokes equations [6], and to incompressible flow problems [7], [26].

Although above stabilization techniques are deeply studied for the convection dominated diffusion equations, their application to optimal control problems governed by convection dominated diffusion equations is not yet intensively studied. This may be due to the fact that stable numerical treatment of the optimality conditions requires stabilization for both the state and the adjoint equation, and it is not straightforward to chose stabilization techniques such that the approaches *first optimize, then discretize* and *first discretize, then optimize* commute. This question for example pops up if one considers the streamline upwind Galerkin method (SUPG) for discretizing the state and the adjoint equation in the optimality system, since this approach seems not to be well suited for the duality techniques frequently used in optimal control. In [3] and [29] stabilized finite element methods for optimal control governed by convection diffusion equations are applied. Both approaches use standard finite element discretization with stabilization based on symmetric penalty terms, where local projections (the so called LPS-method) are used in [3], and edge stabilization (see [5]) in [29]. Then formulating the control problem on the continuous level and then discretizing the optimality conditions appropriately is equivalent to considering the control problem on the discrete level. Hence the question posed above of which concept to apply is made redundant.

In [3] a priori error estimates are proved for both constrained and unconstrained problems, while a priori and a posteriori error estimates are provided in [29].

In [14] the first author proposes the variational discretization concept for optimal control problems with control constraints, which implicitly utilizes the first order optimality conditions and the discretization of the state and adjoint equations for the discretization of the control instead of discretizing the space of admissible controls. The application to the control governed by elliptic equations is discussed, and optimal error estimates are provided.

Here we combine variational discretization and the edge stabilization Galerkin method and apply it to the discretization of optimal control problems governed by convection diffusion equations. We first derive the continuous optimality system, which contains the state equation, the adjoint state equation and the optimality condition, which is given in terms of a variational inequality. Then similar to the standard approaches to optimal control problems governed by elliptic or parabolic partial differential equations (see, e.g., [21]-[24]), we derive the discrete optimal control problem by using the edge stabilization Galerkin method to approximate the state equation, whose optimality system then coincides with that obtained by discretizing the state and adjoint state in the continuous optimality system by finite elements with edge stabilization. The control is not discretized in our approach. For the control u, the state y and the adjoint state p we prove the a priori estimate

$$\| y - y_h \|_{*,\Omega} + \| p - p_h \|_{*,\Omega} + \| u - u_h \|_{0,\Omega} \leq C(h^{3/2} + h\varepsilon^{1/2}),$$

where u_h, y_h, p_h denote their discrete counterparts, and $\| \cdot \|_{*,\Omega}$ is defined in Section 3. We note that this result is of the same quality as those obtained in [3] and [29], but is obtained without structural assumptions like [3, Assumption 2], and also by a different simpler proof technique. Furthermore, we construct a residual type a posteriori error estimator which only contains contributions from the local residuals in the state and the adjoint equation. Contributions from the optimality condition do not appear since the control is not discretized in the variational approach taken here. Finally, numerical examples are presented to illustrate our theoretical results.

The paper is organized as follows: In Section 2, we describe the edge stabilization Galerkin scheme for the constrained optimal control problem governed by convection dominated diffusion equations using variational discretization. In Sections 3 we prove the a priori error estimate, and in Section 4, the a posteriori error estimator is constructed. In the last section, we present two numerical examples to illustrate the theoretical results.

2 Model problem and its variational approximation scheme

In this section we consider the following constrained optimal control problem governed by convection dominated diffusion equations:

$$\min_{u\in K\subset U} J(y,u) \tag{2.1}$$

subject to

$$\begin{aligned} -\varepsilon\triangle y+\vec{b}\cdot\nabla y+ay &= f+u, \quad \text{in} \quad \Omega, \\ y &= 0, \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{2.2}$$

where

$$J(y,u)=\frac{1}{2}\parallel y-y_0\parallel_{0,\Omega}^2+\frac{\alpha}{2}\parallel u\parallel_{0,\Omega}^2,$$

$\alpha>0$ is a constant, $\Omega\subset R^2$ is a bounded domain with Lipschitz boundary $\partial\Omega$, $K\subset U=L^2(\Omega)$ denotes a closed convex set. From here onwards we use

$$K=\{u\in U; u_a\le u\le u_b \ \text{ a.e. in } \ \Omega\},$$

where for simplicity $u_a<u_b$ denote constants. Moreover, $f\in L^2(\Omega)$, $a>0$ is the reaction coefficient, $0<\varepsilon<<1$ is a small positive number, $\vec{b}\in(W^{1,\infty}(\Omega))^2$ is a velocity field. We assume that the following coercivity condition holds:

$$a-\frac{1}{2}\nabla\cdot\vec{b}\ge a_0>0.$$

In this paper we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev space on Ω with a norm $\|\cdot\|_{m,q,\Omega}$ and a semi-norm $|\cdot|_{m,q,\Omega}$. We set $W_0^{m,q}(\Omega)=\{v\in W^{m,q}(\Omega):\ v\mid_{\partial\Omega}=0\}$. For $q=2$, we denote $H^m(\Omega)=W^{m,2}(\Omega)$ and $\|\cdot\|_{m,\Omega}=\|\cdot\|_{m,2,\Omega}$. Especially, we denote the state space by $Y=H_0^1(\Omega)$. The inner product in $L^2(\Omega)$ is indicated by and $(\cdot,\cdot)$. In addition, C denotes a generic constant.

Let us first consider the weak formulation of the state equation. It is well known that the weak formulation of the state equation (2.2) is to find $y(u)\in Y$, such that

$$(\varepsilon\nabla y,\nabla w)+(\vec{b}\cdot\nabla y,w)+(ay,w)=(f+u,w), \quad \forall w\in Y.$$

Let $A(\cdot,\cdot)$ be the bilinear form given by

$$A(y,w)=(\varepsilon\nabla y,\nabla w)+(\vec{b}\cdot\nabla y,w)+(ay,w), \qquad \forall y,w\in Y.$$

We define an energy norm associated with (2.2) via

$$||| y |||_\Omega = \{\varepsilon \parallel \nabla y \parallel_{0,\Omega}^2 + \parallel a_0^{\frac{1}{2}} y \parallel_{0,\Omega}^2\}^{1/2}.$$

It is easy to see that

$$A(y, y) \geq ||| y |||_\Omega^2 . \tag{2.3}$$

Therefore the variational formulation corresponding to (2.1)-(2.2) can be rewritten as

$$\min_{u \in K} J(y, u) \tag{2.4}$$

subject to

$$A(y(u), w) = (f + u, w), \qquad \forall w \in Y. \tag{2.5}$$

Since (2.4)-(2.5) defines a strictly convex optimal control problem it is clear (see e.g. [13] and [20]) that it admits a unique solution (y, u), and that a pair (y, u) is the solution of (2.4)-(2.5) if and only if there is a adjoint state $p \in$ Y, such that (y, p, u) satisfies the following optimality conditions:

$$A(y, w) = (f + u, w), \quad \forall w \in Y, \tag{2.6}$$

$$A(q, p) = (y - y_0, q), \quad \forall q \in Y, \tag{2.7}$$

$$(\alpha u + p, v - u) \geq 0, \quad \forall v \in K. \tag{2.8}$$

It is well known that using the pointwise projection on the admissible set K,

$$P_K : U \longrightarrow K, \quad P_K(v) = \max(u_a, \min(u_b, v)), \tag{2.9}$$

the optimality condition (2.8) can be equivalently expressed as

$$u = P_K\left(-\frac{1}{\alpha}p\right). \tag{2.10}$$

Note that the state equation (2.5) is the convection dominated diffusion equation when ε is very small. It is well known that the standard finite element method can not work well for solving this kind of problems. Stabilized methods should be adopted in order to improve the computational accuracy. The edge stabilization Galerkin scheme (see, e.g., [5]) has been proved to be a efficient scheme for the equation (2.5). In this paper, we use the edge stabilization Galerkin scheme to deal with the state equation and costate equation in the optimality system (2.6)-(2.8).

Let T^h be regular triangulations of Ω, so that $\bar{\Omega} = \cup_{\tau \in T^h} \bar{\tau}$. Let $h = \max_{\tau \in T^h} h_\tau$, where h_τ denotes the diameter of the element τ. Associated with T^h is a finite dimensional subspace W^h of $C(\bar{\Omega})$, such that $\phi|_\tau$ is the polynomial of k-order ($k \geq 1$), $\forall \phi \in W^h$. Set $Y^h = W^h \cap Y$. Then it is easy to see that $Y^h \subset Y = H_0^1(\Omega)$.

To control the advective derivative of the discrete solution sufficiently we introduce a stabilization form S on $Y^h \times Y^h$ (see e.g. [5]) such that

$$S(v_h, w_h) = \sum_{l \in E^h} \int_l \gamma h_l^2 [\vec{n} \cdot \nabla v_h][\vec{n} \cdot \nabla w_h] ds,$$

where $E^h \subset \partial T^h$ denotes the collection of interior edges of the triangles in T^h (∂T^h is the collection of all edges of the triangles in T^h), h_l is the size of the edge l, $[q]_l$ denotes the jump of q across l for $l \in E^h$ such that

$$[q(x)]_{x\in l} = \lim_{s\to 0^+} \Big(q(x + s\vec{n}) - q(x - s\vec{n}) \Big),$$

where $\vec{n}$ is the outward unit normal.

Using the stabilization form defined above, an edge stabilization Galerkin approximation of the optimal control problem (2.4)-(2.5) can be defined as follows

$$\min_{u_h \in K} J(y_h, u_h) \tag{2.11}$$

subject to

$$A(y_h, w_h) + S(y_h, w_h) = (f + u_h, w_h), \qquad \forall w_h \in Y^h. \tag{2.12}$$

As in the continuous case it can be shown that the control problem (2.11)-(2.12) admits a unique solution (y_h, u_h), and that a pair (y_h, u_h) is the solution of (2.11)-(2.12) if and only if there is a unique adjoint state $p_h \in V^h$, such that (y_h, p_h, u_h) satisfies the optimality conditions

$$A(y_h, w_h) + S(y_h, w_h) = (f + u_h, w_h), \quad \forall w_h \in Y^h, \tag{2.13}$$

$$A(q_h, p_h) + S(q_h, p_h) = (y_h - y_0, q_h), \quad \forall q_h \in Y^h, \tag{2.14}$$

$$(\alpha u_h + p_h, v - u_h) \geq 0, \qquad \forall v \in K. \tag{2.15}$$

Concerning (2.15) it should be pointed out that we minimize over the infinite dimensional set K instead of minimizing over a finite-dimensional subset of K. Similar to (2.10), the projection (2.9) allows to rewrite the optimality condition (2.15) as

$$u_h = P_K \left(-\frac{1}{\alpha} p_h \right). \tag{2.16}$$

In general, u_h is not a finite element function corresponding to the mesh T^h, especially on triangles containing the discrete free boundary. This fact requires more care for the construction of the algorithms for computing u_h, see [11] for details.

3 A priori error estimates

In this section, we consider a priori error estimates for the optimal control problem (2.6)-(2.8) and its edge stabilization Galerkin approximation (2.13)-(2.15).

Theorem 3.1. *Let (y, p, u) and (y_h, p_h, u_h) denote the solutions to (2.6)-(2.8) and (2.13)-(2.15), respectively. Assume that $y,\ p \in H^2(\Omega)$. Then we have*

$$\| y - y_h \|_{*,\Omega} + \| p - p_h \|_{*,\Omega} + \| u - u_h \|_{0,\Omega} \leq C(h^{3/2} + h\varepsilon^{1/2}), \tag{3.1}$$

where

$$\| w_h \|_{*,\Omega}^2 = \varepsilon \| \nabla w_h \|_{0,\Omega}^2 + \| a_0^{\frac{1}{2}} w_h \|_{0,\Omega}^2 + \| h^{\frac{1}{2}} \vec{b} \cdot \nabla w_h \|_{0,\Omega}^2 + S(w_h, w_h).$$

Proof. Let $\hat{J}(v) := J(y_h(v), v)$ denote the reduced functional, where for given $v \in U$ the function $y_h(v) \in Y^h$ solves (2.11) with u_h replaced by v. Then straightforward calculation yields $\hat{J}_h'(v) = \alpha v + p_h(v)$, where for given $v \in U$ the function $p_h(v) \in Y^h$ solves

$$A(q_h, p_h(v)) + S(q_h, p_h(v)) = (y_h(v) - y_0, q_h), \quad \forall q_h \in Y^h.$$

Now we test (2.15) with $v = u$, (2.8) with $v = u_h$, and add the resulting inequalities. This implies

$$\alpha \|u - u_h\|_{0,\Omega}^2 \leq (p - p_h, u_h - u) = (p - \tilde{p}_h(u), u_h - u) + (\tilde{p}_h(u) - p_h, u_h - u) := (1) + (2),$$

where for given $v \in U$ the function $\tilde{p}_h(v)$ solves

$$A(q_h, \tilde{p}_h(v)) + S(q_h, \tilde{p}_h(v)) = (y(v) - y_0, q_h), \quad \forall q_h \in Y^h. \tag{3.2}$$

Then

$$|(1)| \leq \|p - \tilde{p}_h(u)\|_{0,\Omega} \|u - u_h\|_{0,\Omega}.$$

Using duality we obtain

$$\begin{aligned}(2) &= A(y_h - y_h(u), \tilde{p}_h(u) - p_h) + S(y_h - y_h(u), \tilde{p}_h(u) - p_h) = \\ &= (y - y_h, y_h - y_h(u)) = -\|y - y_h\|_{0,\Omega}^2 + (y - y_h, y - y_h(u)) \leq \\ &\leq -\frac{1}{2}\|y - y_h\|_{0,\Omega}^2 + \frac{1}{2}\|y - y_h(u)\|_{0,\Omega}^2.\end{aligned}$$

Combining the estimates for (1) and (2) we obtain

$$\alpha \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{0,\Omega}^2 \leq \frac{1}{\alpha}\|p - \tilde{p}_h(u)\|_{0,\Omega}^2 + \|y - y_h(u)\|_{0,\Omega}^2. \tag{3.3}$$

Using the results of [1] and [5], we obtain

$$\| p - \tilde{p}_h \|_{*,\Omega} \leq C(h^{3/2} + h\varepsilon^{1/2}) \|p\|_{2,\Omega}. \tag{3.4}$$

and similarly for y, noting that $y_h(u)$ is the edge stabilization Galerkin solution of y,

$$\| y_h(u) - y \|_{*,\Omega} \leq C(h^{3/2} + h\varepsilon^{1/2}) \|y\|_{2,\Omega}. \tag{3.5}$$

This delivers the intermediate result

$$\|u - u_h\|_{0,\Omega} + \|y - y_h\|_{0,\Omega} \leq C(h^{3/2} + h\varepsilon^{1/2})\{\|y\|_{2,\Omega} + \|p\|_{2,\Omega}\}. \tag{3.6}$$

Finally we estimate $\| y - y_h \|_{*,\Omega}$ and $\| p - p_h \|_{*,\Omega}$. To begin with we recall that $y_h(u)$ is the edge stabilization Galerkin solution of y. By the stability property of $A(\cdot,\cdot) + S(\cdot,\cdot)$ (see,e.g.,[5]) we obtain

$$\| y_h - y_h(u) \|_{*,\Omega} \leq C \| u - u_h \|_{0,\Omega} . \tag{3.7}$$

Similarly,

$$\| p_h - \tilde{p}_h(u) \|_{*,\Omega} \leq C \| y - y_h \|_{0,\Omega}, \tag{3.8}$$

so that the triangle inequality combined with (3.5),(3.7), (3.4),(3.8), and (3.6) gives (3.1). □

Remark 3.2. *A close inspection of the previous proof shows that the result of Theorem 3.1 remains valid for every stabilization S which allows error estimates of the form* (3.5) *and* (3.4), *where the discretization of the adjoint equation is performed according to* (3.2). *Our result therefore immediately applies to the approach taken in [3], since [3, Lemma 5, Lemma 6] are also valid in our setting.*

4 A posteriori error estimates

We now derive residual type a posteriori error estimates in the L^2-norm for problem (2.6)-(2.8) and its edge stabilization Galerkin approximation (2.13)-(2.15). For this purpose we need the following lemma whose proof can be found in [28].

Lemma 4.1. *Let $I_h : H_0^1(\Omega) \to Y^h$ be the interpolation operator of Clement type (see [28]). Then for all $\tau \in T^h$, $l \in E^h$, and $v \in H^1(N(\tau))$ or $v \in H^1(N(l))$, we have*

$$\|v - I_h v\|_{i,\tau} \leq C h_\tau^{k-i} \|v\|_{k,N(\tau)}, \quad 0 \leq i \leq k \leq 1,$$

$$\|v - I_h v\|_{0,l} \leq C h_l^{1/2} \|v\|_{k,N(l)},$$
$$|||I_h v|||_\tau \leq C |||v|||_{N(\tau)},$$

where $E^h \subset \partial T^h$ denotes the collection of interior edges of the triangles in T^h, $N(\tau)$ and $N(l)$ denote the union of all elements that share at least one point with τ and l, and $|||\cdot|||$ is defined in Section 2.

In order to obtain the a posteriori error estimates for $y - y_h$ and $p - p_h$, we introduce the following auxiliary dual problems:

$$\begin{cases} -\varepsilon \triangle \phi_1 - \nabla \cdot (\vec{b}\phi_1) + a\phi_1 = f_1, & \text{in} \quad \Omega, \\ \phi_1 = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{4.1}$$

and

$$\begin{cases} -\varepsilon \triangle \phi_2 + \vec{b} \cdot \nabla \phi_2 + a\phi_2 = f_2, & \text{in} \quad \Omega, \\ \phi_2 = 0, & \text{on} \quad \partial\Omega. \end{cases} \tag{4.2}$$

Noting that $a - \frac{1}{2} \nabla \cdot \vec{b} \geq a_0 > 0$, it is easy to derive the following stability estimates for above auxiliary dual problems.

Lemma 4.2. *Let ϕ_i be the solution of (4.1) or (4.2). For $i = 1$ or 2, we have*

$$\varepsilon \parallel \phi_i \parallel_{1,\Omega}^2 + \parallel \phi_i \parallel_{0,\Omega}^2 \leq C |||\phi_i|||^2 \leq C \parallel f_i \parallel_{0,\Omega}^2 .$$

Now we are in the position to prove the a posteriori error estimate.

Theorem 4.3. *Let (y, p, u) and (y_h, p_h, u_h) denote the solution of (2.6)-(2.8) and (2.13)-(2.15), respectively. Then we have*

$$\|u - u_h\|_{0,\Omega}^2 + \parallel y - y_h \parallel_{0,\Omega}^2 + \parallel p - p_h \parallel_{0,\Omega}^2 \leq \sum_{i=1}^{4} \eta_i^2, \tag{4.3}$$

where

$$\begin{aligned}
\eta_1^2 &= \sum_{\tau\in T^h} \frac{h_\tau^2}{\varepsilon}\int_\tau (f+u_h+\varepsilon\triangle y_h-\vec{b}\cdot\nabla y_h-ay_h)^2,\\
\eta_2^2 &= \sum_{l\in E^h}(\varepsilon+\frac{h_l^2}{\varepsilon})h_l\int_l [\nabla y_h\cdot\vec{n}]^2 ds,\\
\eta_3^2 &= \sum_{\tau\in T^h} \frac{h_\tau^2}{\varepsilon}\int_\tau (y_h-y_0+\varepsilon\Delta p_h+\nabla\cdot(\vec{b}p_h)-ap_h)^2,\\
\eta_4^2 &= \sum_{l\in E^h}(\varepsilon+\frac{h_l^2}{\varepsilon})h_l\int_l [\nabla p_h\cdot\vec{n}]^2 ds,
\end{aligned}$$

$l=\bar{\tau}_l^1\cap\bar{\tau}_l^2$ *denotes the edge of the element,* h_l *its length, and* $[v]_l$ *the jump of* v *over the edge* l*.*

Proof. Set $\hat{J}(u)=J(y(u),u)$ as in the problem (2.1). Then,

$$(\hat{J}'(u),v)=(\alpha u+p,v), \tag{4.4}$$

$$(\hat{J}'(u_h),v)=(\alpha u_h+p(u_h),v), \tag{4.5}$$

where $p(u_h)$ is the solution of the following equations:

$$A(y(u_h),w)=(f+u_h,w),\quad \forall w\in Y, \tag{4.6}$$
$$A(q,p(u_h))=(y(u_h),q),\quad \forall q\in Y. \tag{4.7}$$

It follows from (4.4)-(4.5) that

$$(\hat{J}'(u),u-u_h)-(\hat{J}'(u_h),u-u_h)=\alpha\parallel u-u_h\parallel_{0,\Omega}^2+(p-p(u_h),u-u_h). \tag{4.8}$$

From (4.6)-(4.7) we derive that

$$\begin{aligned}
(p-p(u_h),u-u_h) &= A(y-y(u_h),p-p(u_h)) \\
&= (y-y(u_h),y-y(u_h))\ge 0.
\end{aligned} \tag{4.9}$$

Thus, (4.8) and (4.9) imply that

$$(\hat{J}'(u),u-u_h)-(\hat{J}'(u_h),u-u_h)\ge\alpha\parallel u-u_h\parallel_{0,\Omega}^2. \tag{4.10}$$

Then it follows from (2.8), (2.15) and (4.10) that

$$\begin{aligned}
\alpha\|u-u_h\|_{0,\Omega}^2 &\le (\hat{J}'(u),u-u_h)_U-(\hat{J}'(u_h),u-u_h)\\
&= (\alpha u+p,u-u_h)-(\alpha u_h+p(u_h),u-u_h)\\
&\le (\alpha u_h+p_h,u_h-u)+(p_h-p(u_h),u-u_h)\\
&\le (p_h-p(u_h),u-u_h).
\end{aligned} \tag{4.11}$$

Thus

$$\|u-u_h\|_{0,\Omega}\le C\parallel p_h-p(u_h)\parallel_{0,\Omega}. \tag{4.12}$$

Let $f_1 = y(u_h) - y_h$ in (4.1). Then integration by parts gives

$$\begin{aligned}\| y(u_h) - y_h \|_{0,\Omega}^2 &= (f_1, y(u_h) - y_h)\\ &= (-\varepsilon\triangle\phi_1 - \nabla\cdot(\vec{b}\phi_1) + a\phi_1, y(u_h) - y_h)\\ &= A(y(u_h), \phi_1) - A(y_h, \phi_1).\end{aligned}$$

Note that

$$A(y(u_h), w) = (f + u_h, w), \quad \forall w \in Y,$$

and

$$A(y_h, w_h) + S(y_h, w_h) = (f + u_h, w_h), \quad \forall w_h \in Y^h.$$

Setting $w = \phi_1$ and $w_h = I_h\phi_1$, where I_h is defined in Lemma 4.1, we obtain

$$\begin{aligned}\| y(u_h) - y_h \|_{0,\Omega}^2 &= (f + u_h, \phi_1) - A(y_h, \phi_1 - I_h\phi_1) - A(y_h, I_h\phi_1)\\ &\quad -S(y_h, I_h\phi_1) + S(y_h, I_h\phi_1)\\ &= (f + u_h, \phi_1 - I_h\phi_1) - A(y_h, \phi_1 - I_h\phi_1) + S(y_h, I_h\phi_1)\\ &= \sum_{\tau\in T^h}\int_\tau (f + u_h + \varepsilon\Delta y_h - \vec{b}\cdot\nabla y_h - a y_h, \phi_1 - I_h\phi_1) \qquad (4.13)\\ &\quad + \sum_{l\in E^h}\int_l [\varepsilon\nabla y_h\cdot\vec{n}](I_h\phi_1 - \phi_1)ds + S(y_h, I_h\phi_1)\\ &= I_1 + I_2 + I_3.\end{aligned}$$

Using the approximation properties of the interpolation presented in Lemma 4.1, we have that

$$\begin{aligned}| I_1 | &\leq \sum_{\tau\in T^h} \| f + u_h + \varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - a y_h \|_{0,\tau} \| \phi_1 - I_h\phi_1 \|_{0,\tau}\\ &\leq C\sum_{\tau\in T^h} h_\tau \| f + u_h + \varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - a y_h \|_{0,\tau} \| \nabla\phi_1 \|_{0,N(\tau)} \qquad (4.14)\\ &\leq C(\delta)\sum_{\tau\in T^h}\frac{h_\tau^2}{\varepsilon}\int_\tau (f + u_h + \varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - a y_h)^2 + C\delta\varepsilon \| \phi_1 \|_{1,\Omega}^2,\end{aligned}$$

where δ is an arbitrary positive number. By Lemma 4.2, the first term of (4.13) can be bounded by

$$\begin{aligned}| I_1 | &\leq C(\delta)\sum_{\tau\in T^h}\frac{h_\tau^2}{\varepsilon}\int_\tau (f + u_h + \varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - a y_h)^2 + C\delta \| y(u_h) - y_h \|_{0,\Omega}^2\\ &= C(\delta)\eta_1^2 + C\delta \| y(u_h) - y_h \|_{0,\Omega}^2 . \qquad (4.15)\end{aligned}$$

In a similar way, the second term of (4.13) can be estimated as by

$$\begin{aligned}
I_2 &= \sum_{l\in E^h}\int_l [\varepsilon\nabla y_h\cdot\vec{n}](I_h\phi_1-\phi_1) \\
&\le C\sum_{l\in E^h} h_l^{1/2}\Big(\int_l [\varepsilon\nabla y_h\cdot\vec{n}]^2\Big)^{\frac{1}{2}}\sum_{l\in E^h}\parallel\nabla\phi_1\parallel_{0,N(l)} \qquad (4.16)\\
&\le C(\delta)\sum_{l\in E^h}\varepsilon h_l\int_l [\nabla y_h\cdot\vec{n}]^2 + C\delta\varepsilon\|\phi_1\|_{1,\Omega}^2 \\
&\le C(\delta)\eta_2^2 + C\delta\parallel y(u_h)-y_h\parallel_{0,\Omega}^2 .
\end{aligned}$$

Finally for I_3 we get

$$\begin{aligned}
S(y_h, I_h\phi_1) &= \sum_{l\in E^h}\int_l \gamma h_l^2[\vec{n}\cdot\nabla y_h][\vec{n}\cdot\nabla(I_h\phi_1)]ds \\
&\le C\sum_{l\in E^h} h_l^2\parallel[\vec{n}\cdot\nabla y_h]\parallel_{0,l}\parallel[\nabla(I_h\phi_1)]\parallel_{0,l} .
\end{aligned}$$

Using a well-known inverse inequality, we obtain

$$\parallel[\nabla(I_h\phi_1)]\parallel_{0,l}\le Ch_l^{-\frac{1}{2}}\parallel I_h\phi_1\parallel_{1,\tau_l} . \qquad (4.17)$$

It follows that

$$S(y_h, I_h\phi_1)\le C\sum_{l\in E^h} h_l^{\frac{3}{2}}\parallel[n\cdot\nabla y_h]\parallel_{0,l}\parallel I_h\phi_1\parallel_{1,\tau_l} .$$

Collecting the above estimates and using Lemma 4.1,4.2, we obtain

$$\begin{aligned}
I_3 &\le C(\delta)\sum_{l\in E^h}\frac{h_l^3}{\varepsilon}\int_l [n\cdot\nabla y_h]^2 + C\delta\varepsilon\parallel\phi_1\parallel_{1,\Omega}^2 \qquad (4.18)\\
&\le C(\delta)\eta_2^2 + C\delta\parallel y(u_h)-y_h\parallel_{0,\Omega}^2 .
\end{aligned}$$

Combining (4.13), (4.15), (4.16) and (4.18) we end up with

$$\parallel y(u_h)-y_h\parallel_{0,\Omega}^2\le C(\eta_1^2+\eta_2^2). \qquad (4.19)$$

Similarly, inserting $f_2 = p(u_h) - p_h$ in (4.2), we derive

$$
\begin{aligned}
\| p(u_h) - p_h \|_{0,\Omega}^2 &= (f_2, p(u_h) - p_h) \\
&= (-\varepsilon\triangle\phi_2 + \vec{b}\cdot\nabla\phi_2 + \alpha\phi_2, p(u_h) - p_h) \\
&= A(\phi_2, p(u_h)) - A(\phi_2, p_h) \\
&= (y(u_h) - y_0, \phi_2) - A(\phi_2 - I_h\phi_2, p_h) - A(I_h\phi_2, p_h) \\
&\quad -S(I_h\phi_2, p_h) + S(I_h\phi_2, p_h) \\
&= (y(u_h) - y_h, \phi_2) \\
&\quad +(y_h - y_0 + \varepsilon\Delta p_h + \nabla\cdot(\vec{b}p_h) - ap_h, \phi_2 - I_h\phi_2) \\
&\quad + \sum_{l\cap\partial\Omega=\emptyset}\int_l [\varepsilon\nabla p_h\cdot\vec{n}](I_h\phi_2 - \phi_2)ds + S(p_h, I_h\phi_2) \\
&\le C \| y(u_h) - y_h \|_{0,\Omega} \| \phi_2 \|_{0,\Omega} \\
&\quad + \sum_{\tau\in T^h}\int_\tau (y_h - y_0 + \varepsilon\Delta p_h + \nabla\cdot(\vec{b}p_h) - ap_h)(I_h\phi_2 - \phi_2) \\
&\quad + \sum_{l\cap\partial\Omega=\emptyset}\int_l [\varepsilon\nabla p_h\cdot\vec{n}](I_h\phi_2 - \phi_2)ds \\
&\quad + \sum_{l\in E^h}\int_l \gamma h_l^2 [n\cdot\nabla p_h][n\cdot\nabla(I_h\phi_2)]ds \\
&\le C(\delta)\sum_{\tau\in T^h}\frac{h_\tau^2}{\varepsilon}\int_\tau (y_h - y_0 + \varepsilon\Delta p_h + \nabla\cdot(\vec{b}p_h) - ap_h)^2 \\
&\quad +C(\delta)\sum_{l\in E^h}\varepsilon h_l\int_l [\nabla p_h\cdot\vec{n}]^2 ds + C(\delta)\sum_{l\in E^h}\frac{h_l^3}{\varepsilon}\int_l [n\cdot\nabla p_h]^2 \\
&\quad +C(\delta) \| y(u_h) - y_h \|^2 + C\delta \| p(u_h) - p_h \|_{0,\Omega}^2 .
\end{aligned}
$$

Therefore,

$$
\| p(u_h) - p_h \|_{0,\Omega}^2 \le C(\eta_3^2 + \eta_4^2) + C \| y(u_h) - y_h \|_{0,\Omega}^2 \le C\sum_{i=1}^4 \eta_i^2. \tag{4.20}
$$

Thus (4.12) and (4.20) imply that

$$
\|u - u_h\|_{0,\Omega}^2 \le C\sum_{i=1}^4 \eta_i^2. \tag{4.21}
$$

Moreover, it is easy to see that

$$
\|y - y(u_h)\|_{0,\Omega} \le C\|u - u_h\|_{0,\Omega}, \tag{4.22}
$$

and

$$
\|p - p(u_h)\|_{0,\Omega} \le C\|y - y(u_h)\|_{0,\Omega} \le C\|u - u_h\|_{0,\Omega}. \tag{4.23}
$$

Thus, it can be deduced from (4.19)-(4.23) that

$$
\|y - y_h\|_{0,\Omega}^2 \le C\|y - y(u_h)\|_{0,\Omega}^2 + C\|y(u_h) - y_h\|_{0,\Omega}^2 \le C\sum_{i=1}^4 \eta_i^2, \tag{4.24}
$$

and

$$\|p-p_h\|_{0,\Omega}^2 \le C\|p-p(u_h)\|_{0,\Omega}^2 + C\|p(u_h)-p_h\|_{0,\Omega}^2 \le C\sum_{i=1}^{4}\eta_i^2. \tag{4.25}$$

Then (4.3) follows from (4.21), (4.24) and (4.25). □

Remark 4.4. *When ε is very small, say, $\varepsilon \le Ch^2$, above theorem can be improved. In the new case, we can use the stability estimate $\|\phi_i\|_{0,\Omega} \le \|f_i\|_{0,\Omega}$ instead of $\varepsilon\|\phi_i\|_{1,\Omega}^2 \le \|f_i\|_{0,\Omega}^2$, and replace (4.14), (4.17) by*

$$\begin{aligned} |\ I_1\ | &\le \sum_{\tau\in T^h} \|\ f+u_h+\varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - ay_h\ \|_{0,\tau}\|\ \phi_1 - I_h\phi_1\ \|_{0,\tau} \\ &\le C(\delta)\sum_{\tau\in T^h}\int_\tau (f+u_h+\varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - ay_h)^2 + C\delta\ \|\ \phi_1\ \|_{0,\Omega}^2, \end{aligned}$$

and

$$\|\ [\nabla(I_h\phi_1)]\ \|_{0,l} \le Ch_l^{-\frac{3}{2}}\ \|\ I_h\phi_1\ \|_{0,\tau_l}\ .$$

Then the a posteriori error estimate in Theorem 4.3 can be improved to

$$\|u-u_h\|_{0,\Omega}^2 + \|\ y-y_h\ \|_{0,\Omega}^2 + \|\ p-p_h\ \|_{0,\Omega}^2 \le \sum_{i=1}^{4}\hat{\eta}_i^2,$$

where

$$\begin{aligned} \hat{\eta}_1^2 &= \sum_{\tau\in T^h}\gamma_\tau\int_\tau (f+u_h+\varepsilon\triangle y_h - \vec{b}\cdot\nabla y_h - ay_h)^2, \\ \hat{\eta}_2^2 &= \sum_{l\in E^h}(\varepsilon+\gamma_l)h_l\int_l [\nabla y_h\cdot\vec{n}]^2 ds, \\ \hat{\eta}_3^2 &= \sum_{\tau\in T^h}\gamma_\tau\int_\tau (y_h - y_0 + \varepsilon\Delta p_h + \nabla\cdot(\vec{b}p_h) - ap_h)^2, \\ \hat{\eta}_4^2 &= \sum_{l\in E^h}(\varepsilon+\gamma_l)h_l\int_l [\nabla p_h\cdot\vec{n}]^2 ds, \end{aligned}$$

with

$$\gamma_\tau = \min\{1, \frac{h_\tau^2}{\varepsilon}\}, \quad \gamma_l = \min\{1, \frac{h_l^2}{\varepsilon}\}.$$

5 Numerical examples

In this section we illustrate our theoretical results by numerical examples for the optimization problem

$$\min_{u\in K}\left\{\frac{1}{2}\int_\Omega (y-y_0)^2 + \frac{\alpha}{2}\int_\Omega u^2\right\} \tag{5.1}$$

subject to

$$-\varepsilon\triangle y + \vec{b}\cdot\nabla y + ay = f + Bu \quad \text{in} \quad \Omega, \tag{5.2}$$

where $\Omega = [0,1] \times [0,1]$.

In the numerical simulation, we use the conforming piecewise linear finite element space for the approximation of the state y and the adjoint state p. A projection gradient method (see [14] and [15]) is used to compute the solution of the infinite dimensional optimization problem (2.11), where the projection operator P_K^b in (4.2) of [15] is replaced by P_K defined in (2.9) (see [14] for more details), and $\rho = \frac{1}{\alpha}$. The iteration is stopped if the relative difference of two consecutive iteration is smaller than 1.e-5. The structure of the active sets obtained with variational discretization are depicted in Figure 5.8.

Example 5.1. *Consider problem (5.1)-(5.2) with* $\alpha = 0.1, \vec{b} = (2,3)$, $a = 2$ *and* $\varepsilon = 10^{-3}$. *The admissible set* $K = \{v \in U, v \geq 0\}$. *To examine the convergence properties of the discrete scheme presented in this paper we use the smooth solution*

$$\begin{aligned} y &= 100(1-x_1)^2 x_1^2 x_2(1-2x_2)(1-x_2), \\ p &= 50(1-x_1)^2 x_1^2 x_2(1-2x_2)(1-x_2), \\ u &= \max\{0, -\frac{1}{\alpha}p\}. \end{aligned}$$

to problem (5.1)-(5.2), where the corresponding source functions f *and* y_0 *are obtained by inserting* y, p, u *into the optimality system* (2.6)-(2.7) *and* (2.10).

In this example, the numerical solutions are computed on a series of triangular meshes, which are created from consecutive global refinement of an initial coarse mesh. At each refinement, every triangle is divided into four congruent triangles. Table 5.1 displays the errors of $\| y - y_h \|_{*,\Omega}$, $\| p - p_h \|_{*,\Omega}$ and $\| u - u_h \|_{0,\Omega}$, where Dofs denotes the number of nodes in the meshes. It is shown in Table 5.1 that

$$\begin{aligned} \| y - y_h \|_{*,\Omega} + \| p - p_h \|_{*,\Omega} &= O(h^{\frac{3}{2}}), \\ \| u - u_h \|_{0,\Omega} &= O(h^2), \end{aligned}$$

which are in coincidence with (and better than) our theoretical results on a priori error estimates presented in Section 3.

Table 5.1 Convergence results on a uniform mesh for variational discretization

Dofs	$\|\| y - y_h \|\|_{*,\Omega}$	order	$\|\| p - p_h \|\|_{*,\Omega}$	order	$\|\| u - u_h \|\|_{0,\Omega}$	order
41	5.682e-001		2.332e-001		2.702e-001	
145	2.050e-001	1.47	8.121e-002	1.52	7.748e-002	1.80
545	7.302e-002	1.49	2.883e-002	1.49	1.992e-002	1.96
2113	2.594e-002	1.49	1.030e-002	1.49	5.031e-003	1.99

Let us recall the results obtained by the fully discrete approaches (see, e.g., [3], [29]), where discrete controls are sought in a finite dimensional finite element space $K^h \subset K$. It has been proved there that $\| u - u_h \|_{0,\Omega} = O(h)$ for piecewise constant approximations of the control, and $\| u - u_h \|_{0,\Omega} = O(h^{3/2})$ for piecewise linear approximations of the control. In the piecewise linear case, the convergence order is only $O(h^{\frac{3}{2}})$ instead of the optimal order $O(h^2)$. This is caused by the fact that u may not be smooth near the free boundary even if y and p are smooth there. The numerical results demonstrate above results (see [29]). In order to show the comparison, we present another convergence result in Table 5.2, where u_h is the standard piecewise linear, continuous finite element function, and the projection is chosen as Q_K, where $v_h = Q_K(w)$ for given w is the conforming piecewise linear finite element function with nodal values $v_i = \max\{u_a, \min\{w(x_i), u_b\}\}$. Comparing the results of Table 5.1 with Table 5.2, it turns out that the scheme using variational discretization approximates the control u better than the standard method.

Table 5.2 Convergence results on uniform mesh with conventional discretization

Dofs	$\|\| y - y_h \|\|_{*,\Omega}$	order	$\|\| p - p_h \|\|_{*,\Omega}$	order	$\\| u - u_h \\|_{0,\Omega}$	order
41	5.636e-001		2.341e-001		2.548e-001	
145	2.052e-001	1.46	8.132e-002	1.53	7.962e-002	1.68
545	7.334e-002	1.48	2.883e-002	1.50	2.356e-002	1.76
2113	2.608e-002	1.49	1.030e-002	1.49	7.089e-003	1.73

We present the numerical results for a priori error estimate in Example 5.1. In the next example, we will show the numerical results for a posteriori error estimate.

Example 5.2. *Consider problem (5.1)-(5.2) with $\alpha = 0.1, \vec{b} = (2,3), a = 1, \varepsilon = 10^{-4}$. The exact solutions are taken as*

$$\begin{aligned} y &= \frac{2}{\pi}(arctan(100(-0.5x_1 + x_2 - 0.25))), \\ p &= 16x_1(1-x_1)x_2(1-x_2)(\frac{1}{2} + \frac{1}{\pi}arctan(200(\frac{1}{16} - (x_1 - 1/2)^2 - (x_2 - 1/2)^2)), \\ u &= \max\{-5, \min\{-1, -\frac{1}{\alpha}p\}\}, \end{aligned}$$

and the corresponding source terms f and y_0 again are obtained by inserting u, y, p into the associated optimality system.

It should be pointed that y does not satisfy homogeneous Dirichlet boundary conditions. Clearly the state y develops steep gradients along the line $x_2 - 1/2x_1 - 1/4 = 0$, the adjoint state p develops steep gradients along the circle $\frac{1}{16} - (x_1 - 1/2)^2 - (x_2 - 1/2)^2) = 0$, and the singularity of u is similar to p (see Figures 5.1, 5.4 and 5.8). The purpose of this example is to show that the constructed error estimators are able to detect these areas containing steep gradients. In this example, the adaptive mesh refinement is applied for the finite element approximation. We use $\sum_{i=1}^{2} \eta_i^2$ and $\sum_{i=3}^{4} \eta_i^2$ as the indicators to construct the adaptive finite element mesh T^h for the state y and the adjoint state p, respectively. Figure 5.1 and Figure 5.4 show surfaces of the state y and adjoint state p. Numerical solutions y_h and p_h are presented in Figure 5.3 and Figure 5.6. Figure 5.2 and Figure 5.5 show adaptive meshes obtained with the indicators $\sum_{i=1}^{2} \eta_i^2$ and $\sum_{i=3}^{4} \eta_i^2$, respectively, while Figure 5.7 presents the adaptive mesh obtained with the whole estimator $\sum_{i=1}^{4} \eta_i^2$. It is shown that the y-mesh and p-mesh adapt the areas with steep gradients very well. Furthermore, Table 5.3 presents the comparison of the errors of the state y and the costate p on the uniform mesh and the adaptive mesh. The error of $y - y_h, p - p_h$ and $u - u_h$ on the adaptive mesh with 1174 nodes is similar to the error on the uniform mesh with 2113 nodes. It can be deduced from Table 5.3 that substantial computing work can be saved by using the adaptive finite element method.

Figure 5.8 shows the boarders of the active sets of the continuous solution (red) together with those obtained by variational discretization (blue) and conventional discretization of the controls with piecewise linear, continuous finite elements (cyan). The zoom clearly shows that the boarder of the active set in the case of variational discretization crosses element edges and is not restricted to finite element edges as is the boarder of the active set in the conventional approach.

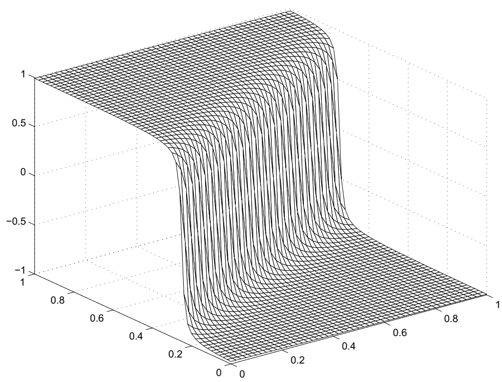

Figure 5.1. The surface of the state y

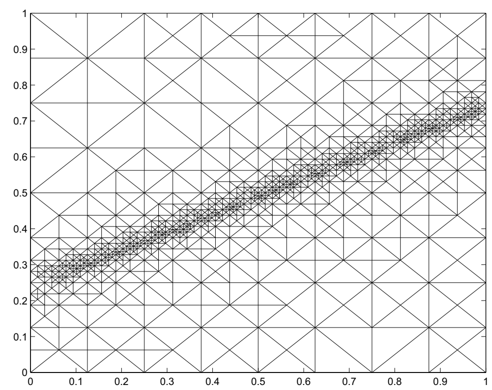

Figure 5.2. The adaptive mesh for the state y

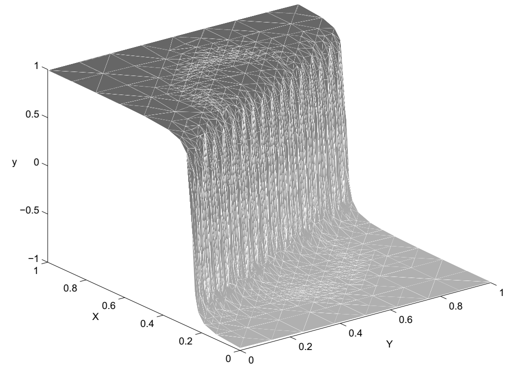

Figure 5.3. The numerical solution y_h

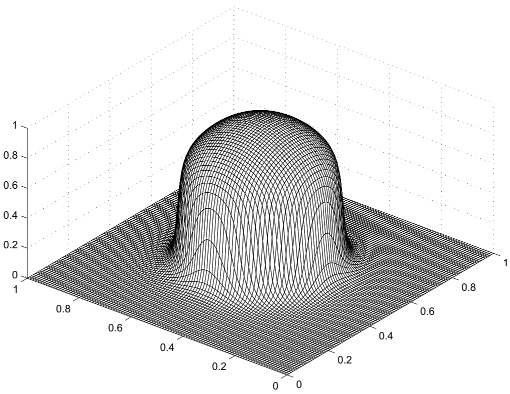

Figure 5.4. The surface of the adjoint state p

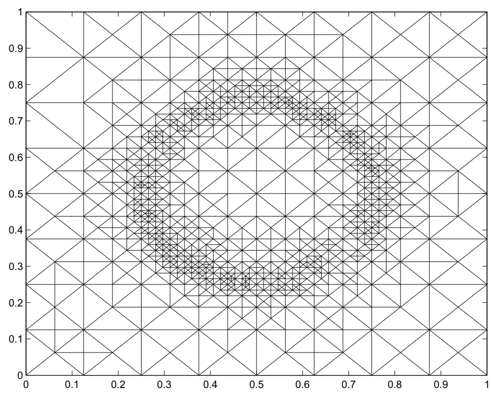

Figure 5.5. The adaptive mesh for the adjoint state p

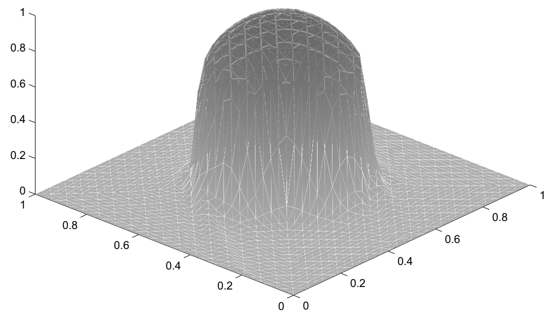

Figure 5.6. The numerical solution p_h

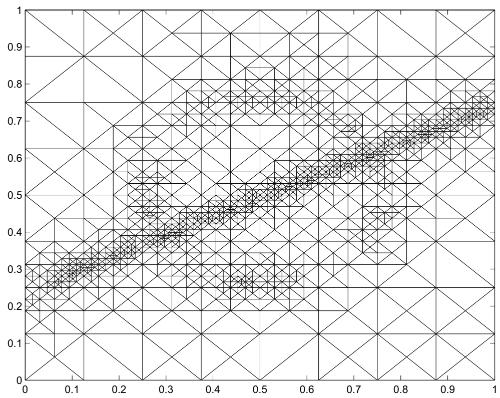

Figure 5.7. The adaptive mesh for both y_h and p_h

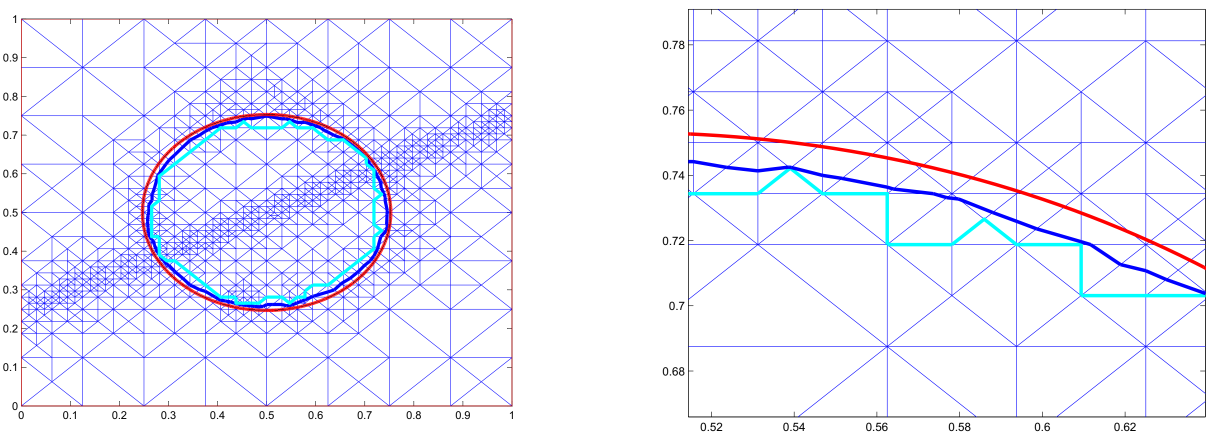

Figure 5.8. Left: The blue line depicts the boarder of the active set of the continuous solution, the black lines depict the boarder of the active set when using variational discretization with piecewise linear, continuous states, and the cyan line depicts the boarder of the active set obtained by using piecewise linear, continuous controls, Right: Zoom

Table 5.3 Comparison of the error of y, p and u on uniform and adaptive meshes

	uniform mesh, nodes=2113	adaptive mesh, nodes=1174
$\parallel y - y_h \parallel_{0,\Omega}$	3.259e-002	1.492e-002
$\parallel p - p_h \parallel_{0,\Omega}$	1.100e-002	1.589e-002
$\parallel y - y_h \parallel_{*,\Omega}$	6.191e-002	3.891e-002
$\parallel p - p_h \parallel_{*,\Omega}$	4.213e-002	4.593e-002
$\parallel u - u_h \parallel_{0,\Omega}$	7.245e-002	9.854e-002

References

[1] L. El Alaoui, A. Ern and E. Burman, A priori and a posteriori analysis of nonconforming finite elements with face penalty for advection–diffusion equations, IMA J. Numer. Anal., 27(1), 151-171(2006).

[2] R. Bartlett, M. Heinkenschloss, D. Ridzal and B. Van Bloemen Waanders, Domain decomposition methods for advection dominated linear-quadratic elliptic optimal control problems, Technical Report SAND 2005-2895, Sandia National Laboratories(2005).

[3] R. Becker and B.Vexler, Optimal control of the convection-diffusion equation using stabilized Finite Element Methods, Numer. Math., 106(3), 349-367(2007).

[4] F. Brezzi and A. Russo, Choosing bubbles for advection-diffusion problems, Math. Models Meth. Appl. Sci., 4, 571-587(1994).

[5] E. Burman and P. Hansbo, Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems, Comput. Methods Appl. Mech. Engrg., 193, 1437-1453(2004).

[6] E. Burman and P. Hansbo, Edge stabilization for Galerkin approximations of the generalized Stokes' problem: A continuous interior penalty method, Comput. Methods Appl. Mech. Engrg, 195(19-22), 2393-2410(2006).

[7] E. Burman and P. Hansbo, A stabilized non-conforming finite element method for incompressible flow, Comput. Methods Appl. Mech. Engrg., 195(23-24), 2881–2899(2006).

[8] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam(1978).

[9] Ph. Clement, Approximation by finite element functions using local regularization, RAIRO Anal. Numer., 2, 77-84(1975).

[10] S. Scott Collis and M. Heinkenschloss, Analysis of the Streamline Upwind/Petrov Galerkin Method applied to the solution of optimal control problems, CAAM TR02-01, March(2002).

[11] J. Douglas Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods(R. Glowinski and J. L. Lions(Eds)), Computing Methods in Applied Sciences, Springer-Verlag, Berlin(1976).

[12] S. Fucik, O. John and A. Kufner, Function Spaces, Nordhoff, Leyden, The Netherland(1977).

[13] A.V. Fursikov, Optimal Control of Distributed Systems. Theory and Applications, American Mathematical Society Providence, Rhode Island(2000).

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

[15] Y.Q. Huang, R. Li, W.B. Liu and N. Yan, Efficient discretization to finite element approximation of constrained optimal control problems, to appear.

[16] T. J. R. Hughes and A. Rooks, Streamline upwind/Petrov Galerkin formulations for the convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg., 54, 199-259(1982).

[17] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge Univ. Press, Cambridge(1987).

[18] C. Johnson and J. Pitkränta, An analysis of the discontinuous Galerkin method for scalar hyperbolic equation, Math. Comp., 46, 1-26(1986).

[19] R. Li, W.B. Liu, H.P. Ma and T. Tang, Adaptive Finite Element approximation for distributed elliptic optimal control problems, SIAM J. Control Optim., 41, 1321-1349(2002).

[20] J. L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin(1971).

[21] W. Liu and N. Yan, A posteriori error estimates for optimal boundary control, SIAM J. Numer. Anal., 39, 73-99(2001).

[22] W. Liu and N. Yan, A posteriori error estimates for distributed convex optimal control problems, Advances in Computational Mathematics., 15, 285-309(2001).

[23] W. Liu and N. Yan, A posteriori error estimates for control problems governed by Stokes equations, SIAM J. Numer. Anal., 40, 1850-1869(2002).

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

[25] U. Navert, A finite element method for convection diffusion problems, Ph.D. thesis, Chalmers Inst. of Tech. (1982).

[26] A. Ouazzi and S. Turek, Unified edge-oriented stabilization of nonconforming finite element methods for incompressible flow problems, J. Numer. Math., (2005).

[27] D. Parra-Guevara and Y.N. Skiba, Elements of the mathematical modeling in the control of pollutants emissions, Ecological Modelling, 167, 263-275(2003).

[28] L. R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp., 54(1990), pp. 483-493.

[29] N. Yan and Z. Zhou, A priori and a posteriori error analysis of edge stabilization Galerkin method for the optimal control problem governed by convection dominated diffusion equation, accepted by Journal of Computational and Applied Mathematics.

[30] J. Zhu and Q. Zeng, A mathematical theoretical frame for control of air pollution, Science in China, Series D, 32, 864-870(2002).