Hamburger Beiträge
zur Angewandten Mathematik

A mixed finite element approximation for optimal control problems with convection-diffusion equations

Michael Hinze, Ningning Yan, and Zhaojie Zhou

Nr. 2011-04
January 2011

A mixed finite element approximation for optimal control problems with convection–diffusion equations

Michael Hinze[†] Ningning Yan[*] and Zhaojie Zhou[†◇]

Abstract

In this paper, we investigate a mixed finite element approximation of convection diffusion optimal control problem without constraints. Under some assumptions about regularity and mesh we prove a second order convergence results for state variable y, adjoint state variable p and control variable u with piecewise linear polynomial approximations. Finally, numerical examples are presented to verify the theoretical findings.

Key words. Optimal control, convection diffusion, mixed finite element, error estimate.
Subject Classification: 65N30.

1 Introduction

Optimal control of convection diffusion equations plays an important role in many practical applications, such as air pollution([7]) and hydraulic pollution([9]) control problems. Taking air pollution control problem as an example, in this process we aim at controlling the emissions of pollutant in order to keep the concentration of pollutant below a certain level over an observation area. In this case the control function is a source term, while the pollutant concentration is described by convection diffusion equations.

Recently, extensive researches have been carried out in this field. In [1], Becker and Vexler apply local projection stabilization method to solve numerically optimal control of convection diffusion equations. A priori error estimates are proved for both unconstrained and constrained problems in [1], while a priori error estimate for a edge stabilization finite element approximation of convection diffusion optimal control problems with constraints is obtained by the second and the third authors in [4]. In [3] and [6] a priori error estimates for streamline diffusion finite element approximations of optimal control problem of convection diffusion equations without constraints and with constraints are investigated, respectively. In [12] the authors discuss variational discretization [8] for optimal control governed by convection dominated diffusion equations. For more numerical methods the reader may refer to the references cited therein.

So far the error estimates in the work cited above only contain $O(h^{3/2})$ convergence results with piecewise linear polynomial approximation.

In the present work, we propose a new mixed finite element method to solve numerically the convection diffusion optimal control problems with and without constraints on the control.

Firstly, for one-dimensional unconstrained control problems we obtain a priori error estimate for state y, adjoint state p and control u as follows

$$\|y-y_h\|+\|p-p_h\|+\|u-u_h\| \leqslant Ch^2,$$

[†] Department of Mathematics, University of Hamburg, 20146 Hamburg, Germany.
[*]Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100190 Beijing, China.
[◇]School of Mathematics Sciences, Shandong Normal University, 250014 Ji'nan, China.

where y_h, p_h and u_h denote the discrete counterparts. For two-dimensional unconstrained control problems we derive a priori error estimate for state y, adjoint state p and control u of the form

$$\|y-y_h\|+\|p-p_h\|+\|u-u_h\| \leqslant \begin{cases} Ch, & \text{on a general mesh,} \\ Ch^2, & \text{on a uniform rectangular mesh.} \end{cases}$$

Secondly, for two-dimensional constrained control problems we utilize the mixed finite element method proposed in this paper to approximate the adjoint state variable, while a edge stabilization Galerkin method is applied to solve the state variable. We derive a priori error estimates for the adjoint p and the control u

$$\|p-p_h\|+\|u-u_h\| \leqslant \begin{cases} Ch, & \text{on a general mesh,} \\ Ch^2, & \text{on a uniform rectangular mesh.} \end{cases}$$

For the state y we obtain the following estimates

$$\| y-y_h \|_* \leqslant \begin{cases} C(h+h^{\frac{3}{2}}+h\varepsilon^{\frac{1}{2}}), & \text{on a general mesh,} \\ C(h^2+h^{\frac{3}{2}}+h\varepsilon^{\frac{1}{2}}), & \text{on a uniform rectangular mesh.} \end{cases}$$

Finally, numerical examples are carried out to verify the theoretical findings.

The paper is organized as follows. In section 2 we prove a priori error estimate for state y, adjoint state p and control u for the unconstrained control problem. In section 3 we consider the constrained control problems and derive the corresponding error estimates. Finally, in section 4 numerical examples are presented to illustrate our analytical findings.

2 Optimal control problem without constraints

2.1 One dimensional Case

In this section we consider the following one-dimensional convection diffusion optimal control problems without constraints;

$$\min_{u\in U} J(y,u)=\frac{1}{2}\int_\Omega (y-z)^2dx+\frac{\alpha}{2}\int_\Omega u^2dx \tag{2.1}$$

subject to

$$-\varepsilon y''+y' = f+u, \quad \text{in } \Omega, \tag{2.2}$$
$$y = 0, \quad \text{on } \partial\Omega. \tag{2.3}$$

Here $\Omega=(a,b)\subset R$ is a one-dimensional bounded domain, and α is a positive constant. f and z are given functions. $U=L^2(\Omega)$, and $\varepsilon>0$ denotes the diffusion coefficient.

Now we are in a position to approximate the convection diffusion optimal control problems (2.1)-(2.3). Here we propose a mixed finite element method to solve numerically the optimal control problems (2.1)-(2.3).

By standard argument we derive the following first order optimality conditions for optimal control problems (2.1)-(2.3);

$$-\varepsilon y''+y' = f+u, \quad x\in\Omega, \tag{2.4}$$
$$-\varepsilon p''-p' = y-z, \quad x\in\Omega, \tag{2.5}$$
$$u+\frac{p}{\alpha} = 0, \quad x\in\Omega, \tag{2.6}$$

$$y = 0,\ p \quad = 0, \qquad x \in \partial\Omega. \tag{2.7}$$

Substituting (2.6) into (2.4) leads to

$$-\varepsilon y'' + y' \quad = \quad f - \frac{p}{\alpha},$$

which implies

$$p = \alpha f + \alpha\varepsilon y'' - \alpha y'.$$

Inserting the expression of p into the adjoint state equation (2.5) yields

$$-\alpha\varepsilon^2 y^{(4)} + \alpha y'' - y \quad = \quad \alpha\varepsilon f'' + \alpha f' - z \triangleq -F_1, \tag{2.8}$$

where $y^{(4)} = \frac{d^4y}{dx^4}$.

To define mixed finite element discrete scheme we introduce $w = \sqrt{\alpha}\varepsilon y''$ and split (2.8) into the following systems

$$w - \sqrt{\alpha}\varepsilon y'' \quad = \quad 0, \tag{2.9}$$

$$\sqrt{\alpha}\varepsilon w'' - \alpha y'' + y \quad = \quad F_1. \tag{2.10}$$

For the boundary condition we have

$$y = 0, \quad y' = f + \frac{w}{\sqrt{\alpha}}, \quad \text{on} \ \ \partial\Omega. \tag{2.11}$$

The latter equation implies

$$p = \alpha f + \sqrt{\alpha} w - \alpha y' = 0, \quad \text{on} \ \ \partial\Omega.$$

For the discretization of problems (2.9)-(2.10) we consider a shape regular mesh $\mathcal{T}_h$ that partition the computational domain Ω into intervals $I_i = [x_i, x_{i+1}]$. Let h_i denote the length of interval I_i and $h = \max_i\{h_i\}$. Moreover, we define

$$(v, \lambda) = \int_a^b v \cdot \lambda dx, \quad < v, \omega >= (v \cdot \lambda)|_a^b. \tag{2.12}$$

Multiplying (2.9) and (2.10) by $\psi \in H^1(\Omega)$ and $\varphi \in H_0^1(\Omega)$ leads to the weak formulation of (2.9) and (2.10);

$$-\sqrt{\alpha}\varepsilon(w', \varphi') + (\alpha y', \varphi') + (y, \varphi) \quad = \quad (F_1, \varphi), \tag{2.13}$$

$$(w, \psi) + \sqrt{\alpha}\varepsilon(y', \psi') - < \sqrt{\alpha}\varepsilon y', \psi > \quad = \quad 0. \tag{2.14}$$

Note that $y' = f + \frac{w}{\sqrt{\alpha}}$ on the boundary. We can rewrite (2.13)-(2.14) as follows

$$-\sqrt{\alpha}\varepsilon(w', \varphi') + (\alpha y', \varphi') + (y, \varphi) \quad = \quad (F_1, \varphi), \qquad \forall \varphi \in H_0^1(\Omega), \tag{2.15}$$

$$(w, \psi) + \sqrt{\alpha}\varepsilon(y', \psi') - < \varepsilon w, \psi > \quad = \quad < \sqrt{\alpha}\varepsilon f, \psi >, \quad \forall \psi \in H^1(\Omega). \tag{2.16}$$

Let $V_h \subset H^1(\Omega)$ be a finite element space consisting of piecewise linear polynomials. Set $V_h^0 = V_h \cap H_0^1(\Omega)$.

Then the mixed finite element approximation of (2.9)-(2.10) is given by

$$-\sqrt{\alpha}\varepsilon(w_h', \varphi_h') + (\alpha y_h', \varphi_h') + (y_h, \varphi_h) \quad = \quad (F_1, \varphi_h), \qquad \forall \varphi_h \in V_h^0, \tag{2.17}$$

$$(w_h, \psi_h) + \sqrt{\alpha}\varepsilon(y_h', \psi_h') - < \varepsilon w_h, \psi_h > \quad = \quad < \sqrt{\alpha}\varepsilon f, \psi_h >, \quad \forall \psi_h \in V_h. \tag{2.18}$$

In the following, let us investigate the error estimates of $y - y_h$ and $w - w_h$.

Firstly, Testing (2.15) with φ_h, (2.16) with ψ_h and subtracting (2.17)-(2.18) from the resulting equations leads to the following error equations:

$$
\begin{aligned}
-\sqrt{\alpha}\varepsilon(w' - w_h', \varphi_h') + (\alpha(y' - y_h'), \varphi_h') + (y - y_h, \varphi_h) &= 0, \quad \forall \varphi_h \in V_h^0, && (2.19)\\
(w - w_h, \psi_h) + \sqrt{\alpha}\varepsilon(y' - y_h', \psi_h') - < \varepsilon(w - w_h), \psi_h > &= 0, \quad \forall \psi_h \in V_h. && (2.20)
\end{aligned}
$$

Let y_I and w_I denote the Lagrange interpolation of y and w, respectively. Then we can decompose the errors $y - y_h$ and $w - w_h$ as

$$
\begin{aligned}
y - y_h &= y - y_I - (y_h - y_I) = \eta_y - \xi_y,\\
w - w_h &= w - w_I - (w_h - w_I) = \eta_w - \xi_w.
\end{aligned}
$$

Then the error equations (2.19)-(2.20) can be expressed as follows

$$
\begin{aligned}
-\sqrt{\alpha}\varepsilon(\xi_w', \varphi_h') + (\alpha\xi_y', \varphi_h') + (\xi_y, \varphi_h) &= -\sqrt{\alpha}\varepsilon(\eta_w', \varphi_h') + (\alpha\eta_y', \varphi_h') + (\eta_y, \varphi_h), && (2.21)\\
(\xi_w, \psi_h) + \sqrt{\alpha}\varepsilon(\xi_y', \psi_h') &= (\eta_w, \psi_h) + \sqrt{\alpha}\varepsilon(\eta_y', \psi_h') - < \varepsilon(\eta_w - \xi_w), \psi_h > . && (2.22)
\end{aligned}
$$

Choosing $\varphi_h = \xi_y$ and $\psi_h = \xi_w$ in (2.21) and (2.22), respectively, and adding the resulting equations together yields

$$
\begin{aligned}
\|\xi_y\|^2 + \|\xi_w\|^2 + \alpha\|\xi_y'\|^2 &= -\sqrt{\alpha}\varepsilon(\eta_w', \xi_y') + (\alpha\eta_y', \xi_y') + (\eta_y, \xi_y),\\
&+ (\eta_w, \xi_w) + \sqrt{\alpha}\varepsilon(\eta_y', \xi_w') - < \varepsilon(\eta_w - \xi_w), \xi_w >,\\
&= \sum_{i=1}^{6} A_i. && (2.23)
\end{aligned}
$$

Next let us discuss the estimates of the related terms on the right hand of (2.23), i.e., $A_1 \sim A_6$. Using the Cauchy-Schwartz inequality we obtain

$$
A_3 \leqslant \frac{1}{2}\|\eta_y\|^2 + \frac{1}{2}\|\xi_y\|^2,
$$

and

$$
A_4 \leqslant \frac{1}{2}\|\eta_w\|^2 + \frac{1}{2}\|\xi_w\|^2.
$$

Note that $\xi_y \in V_h^0$ and $\xi_w \in V_h$. Together with the definition of Lagrange interpolation we have

$$
A_1 = 0, \quad A_2 = 0, \quad A_5 = 0.
$$

For the last term A_6 we have

$$
\begin{aligned}
A_6 &= - < \varepsilon(\eta_w - \xi_w), \xi_w >\\
&= < \varepsilon\xi_w, \xi_w >\\
&= \int_\Omega \varepsilon(\xi_w^2)' dx\\
&= 2\int_\Omega \varepsilon\xi_w\xi_w' dx\\
&\leqslant 2\varepsilon\|\xi_w\| \cdot \|\xi_w'\|\\
&\leqslant 2\varepsilon\frac{\sqrt{12}}{\underline{h}}\|\xi_w\|^2, && (2.24)
\end{aligned}
$$

where we have used the inverse estimate

$$|\psi_h|_1^2 \leqslant \frac{12}{\underline{h}^2}\|\psi_h\|^2,$$

with $\underline{h} = \min_i\{h_i\}$. Inserting the estimates of $A_1 \sim A_6$ into (2.23) we have

$$\frac{1}{2}\|\xi_y\|^2 + \frac{1}{2}\|\xi_w\|^2 + \alpha\|\xi_y'\|^2 \leqslant \frac{1}{2}\|\eta_y\|^2 + \frac{1}{2}\|\eta_w\|^2 + 2\varepsilon\frac{\sqrt{12}}{\underline{h}}\|\xi_w\|^2.$$

Assume that $2\varepsilon\frac{\sqrt{12}}{\underline{h}} \leqslant 1/2 - \delta$ for a small positive constant δ, i.e., $\varepsilon < \frac{\underline{h}}{8\sqrt{3}}$. Then we have

$$\begin{aligned}\frac{1}{2}\|\xi_y\|^2 + \delta\|\xi_w\|^2 + \alpha\|\xi_y'\|^2 &\leqslant \frac{1}{2}\|\eta_y\|^2 + \frac{1}{2}\|\eta_w\|^2 \\ &\leqslant C(\|y\|_2^2 + \|w\|_2^2)h^4. \end{aligned} \tag{2.25}$$

Utilizing (2.25) and the estimate of Lagrange interpolation we deduce the final error estimates, which can be stated as

Theorem 2.1. *Let (y, w) and (y_h, w_h) be the solutions of (2.15)-(2.16) and (2.17)-(2.18), respectively. Assume that $\varepsilon < \frac{\underline{h}}{8\sqrt{3}}$. Then, there exists a constant C independent of h and ε such that*

$$\begin{aligned}\|y - y_h\| + \|w - w_h\| &\leqslant C(\|y\|_2 + \|w\|_2)h^2, \\ \|y' - y_h'\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + Ch\|y\|_2.\end{aligned}$$

Proceeding similarly as the treatment to state variable y we deduce the following equation for adjoint state variable p

$$-\alpha\varepsilon^2 p^{(4)} + \alpha p'' - p = -\alpha\varepsilon z'' + \alpha z' - \alpha f \triangleq -G_1. \tag{2.26}$$

By introducing $\vartheta = \sqrt{\alpha}\varepsilon p''$ we can rewrite (2.26) as the following systems

$$\vartheta - \sqrt{\alpha}\varepsilon p'' = 0, \tag{2.27}$$

$$\sqrt{\alpha}\varepsilon\vartheta'' - \alpha p'' + p = G_1. \tag{2.28}$$

On the boundary we have

$$p = 0, \quad p' = z - \frac{\vartheta}{\sqrt{\alpha}}, \quad \text{on } \partial\Omega. \tag{2.29}$$

The latter equation implies

$$y = 0, \quad \text{on } \partial\Omega.$$

Then the weak formulation for adjoint state variable state p is given by

$$-\sqrt{\alpha}\varepsilon(\vartheta', \varphi') + (\alpha p', \varphi') + (p, \varphi) = (G_1, \varphi), \qquad \forall\varphi \in H_0^1(\Omega), \tag{2.30}$$

$$(\vartheta, \psi) + \sqrt{\alpha}\varepsilon(p', \psi') + < \varepsilon\vartheta, \psi > = < \sqrt{\alpha}\varepsilon z, \psi >, \quad \forall\psi \in H^1(\Omega). \tag{2.31}$$

The mixed finite element approximation of p is to find $(p_h, \vartheta_h) \in V_h \times V_h^0$ satisfying

$$-\sqrt{\alpha}\varepsilon(\vartheta_h', \varphi_h') + (\alpha p_h', \varphi_h') + (p_h, \varphi_h) = (G_1, \varphi_h), \qquad \forall\varphi_h \in V_h^0, \tag{2.32}$$

$$(\vartheta_h, \psi_h) + \sqrt{\alpha}\varepsilon(p_h', \psi_h') + < \varepsilon\vartheta_h, \psi_h > = < \sqrt{\alpha}\varepsilon z, \psi_h >, \; \forall\psi_h \in V_h. \tag{2.33}$$

For the adjoint variable p, ϑ and their approximations p_h, ϑ_h we have the following results.

Theorem 2.2. *Let (p, ϑ) and (p_h, ϑ_h) be the solutions of (2.30)-(2.31) and (2.32)-(2.33), respectively. Assume that $\varepsilon < \frac{h}{8\sqrt{3}}$. Then, there exists a constant C independent of h and ε such that*

$$\begin{aligned} \|p - p_h\| + \|\vartheta - \vartheta_h\| &\leqslant Ch^2(\|p\|_2 + \|\vartheta\|_2), \\ \|p' - p'_h\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_2 + \|\vartheta\|_2) + Ch\|p\|_2. \end{aligned}$$

Proof. The proof of this theorem is along the lines of the proof of Theorem 2.1 and therefore omitted. □

2.2 Two dimensional Case

In this section we consider the following two dimensional convection diffusion optimal control problems without constraints;

$$\min_{u \in U} J(y, u) = \frac{1}{2}\int_\Omega (y - z)^2 dx + \frac{\alpha}{2}\int_\Omega u^2 dx \tag{2.34}$$

subject to

$$-\varepsilon\Delta y + \boldsymbol{\beta} \cdot \nabla y = f + u, \quad \text{in } \Omega, \tag{2.35}$$

$$y = 0, \quad \text{on } \partial\Omega. \tag{2.36}$$

Here $\Omega \subset R^2$ is a bounded domain, and α is a positive constant. f and z are given functions. $U = L^2(\Omega)$. $\boldsymbol{\beta}$ is constant vector, and $\varepsilon > 0$ denotes the diffusion coefficient.

Again by standard argument we have the following first order optimality conditions for optimal control problems (2.34)-(2.36);

$$-\varepsilon\Delta y + \boldsymbol{\beta} \cdot \nabla y = f + u, \quad x \in \Omega, \tag{2.37}$$

$$-\varepsilon\Delta p - \boldsymbol{\beta} \cdot \nabla p = y - z, \quad x \in \Omega, \tag{2.38}$$

$$u = -\frac{p}{\alpha}, \quad x \in \Omega, \tag{2.39}$$

$$y = 0, \; p = 0, \quad x \in \partial\Omega. \tag{2.40}$$

Proceeding similarly as in the one-dimensional problem we apply a mixed finite element method to solve numerically the optimal control problems (2.34)-(2.36). Substituting (2.39) into (2.37) leads to

$$-\varepsilon\Delta y + \boldsymbol{\beta} \cdot \nabla y = f - \frac{p}{\alpha},$$

which implies in turn

$$p = \alpha f + \alpha\varepsilon\Delta y - \alpha\boldsymbol{\beta} \cdot \nabla y.$$

Inserting the above equality into the adjoint state equation (2.38) yields

$$-\alpha\varepsilon^2\Delta^2 y + \alpha\boldsymbol{\beta} \cdot \nabla(\boldsymbol{\beta} \cdot \nabla y) - y = \alpha\varepsilon\Delta f + \alpha\boldsymbol{\beta} \cdot \nabla f - z \triangleq -F_2. \tag{2.41}$$

Similarly as in one-dimensional problem we introduce $w = \sqrt{\alpha}\varepsilon\Delta y$ and decompose (2.41) into the following systems

$$w - \sqrt{\alpha}\varepsilon\Delta y = 0, \tag{2.42}$$

$$\sqrt{\alpha}\varepsilon\Delta w - \alpha\boldsymbol{\beta} \cdot \nabla(\boldsymbol{\beta} \cdot \nabla y) + y = F_2. \tag{2.43}$$

For the boundary condition we have

$$y = 0, \quad \boldsymbol{\beta} \cdot \nabla y = f + \frac{w}{\sqrt{\alpha}}, \quad \text{on} \;\; \partial\Omega. \tag{2.44}$$

The latter equation implies

$$p = \alpha f + \sqrt{\alpha} w - \alpha \boldsymbol{\beta} \cdot \nabla y = 0, \quad \text{on} \;\; \partial\Omega.$$

Let $\mathbf{n}$ denote the unit outward normal vector and $\boldsymbol{\tau}$ denote the unit tangent vector. Then $\boldsymbol{\beta}$ can be expressed as

$$\boldsymbol{\beta} = (\boldsymbol{\beta} \cdot \mathbf{n})\mathbf{n} + (\boldsymbol{\beta} \cdot \boldsymbol{\tau})\boldsymbol{\tau}.$$

Set

$$\Gamma_1 = \{x \in \partial\Omega; \boldsymbol{\beta} \cdot \mathbf{n}(x) = 0\},$$

and

$$\Gamma_2 = \partial\Omega \backslash \Gamma_1.$$

Further we have

$$\boldsymbol{\beta} \cdot \nabla y = (\boldsymbol{\beta} \cdot \boldsymbol{\tau})\boldsymbol{\tau} \cdot \nabla y = (\boldsymbol{\beta} \cdot \boldsymbol{\tau}) \frac{\partial y}{\partial \boldsymbol{\tau}} = 0, \quad \text{on} \;\; \Gamma_1,$$

and

$$\boldsymbol{\beta} \cdot \nabla y = (\boldsymbol{\beta} \cdot \mathbf{n})\mathbf{n} \cdot \nabla y + (\boldsymbol{\beta} \cdot \boldsymbol{\tau})\boldsymbol{\tau} \cdot \nabla y = (\boldsymbol{\beta} \cdot \mathbf{n}) \frac{\partial y}{\partial \mathbf{n}} + (\boldsymbol{\beta} \cdot \boldsymbol{\tau}) \frac{\partial y}{\partial \boldsymbol{\tau}} = (\boldsymbol{\beta} \cdot \mathbf{n}) \frac{\partial y}{\partial \mathbf{n}}, \quad \text{on} \;\; \Gamma_2.$$

Summarizing, the second equality of (3.11) can be rewritten as

$$w = -\sqrt{\alpha} f, \quad \text{on} \;\; \Gamma_1, \tag{2.45}$$

and

$$\frac{\partial y}{\partial \mathbf{n}} = \frac{1}{(\boldsymbol{\beta} \cdot \mathbf{n})}(f + \frac{w}{\sqrt{\alpha}}), \quad \text{on} \;\; \Gamma_2. \tag{2.46}$$

In this paper, we assume that $\min_{x \in \Gamma_2} (\boldsymbol{\beta} \cdot \mathbf{n}) \geqslant c > 0$.

For the discretization of problems (2.42)-(2.43) we consider a shape regular mesh $\mathcal{T}_h = \{K\}$ which subdivides the computational domain Ω into triangles or parallelograms. The diameter of an element K and the length of an edge e are denoted by h_K and h_e, respectively. Let $h = \max_K \{h_K\}$, $\underline{h} = \min_K \{h_K\}$ and $\hat{H}_0^1(\Omega) = \{v \in H^1(\Omega); v|_{\Gamma_1} = 0\}$. Moreover, we define

$$(v, \omega) = \int_\Omega v \cdot \omega dx, \quad < v, \omega >= \int_\Gamma v \cdot \omega ds, \tag{2.47}$$

where $\Gamma = \partial\Omega$ or a part of $\partial\Omega$.

Multiplying (2.42) and (2.43) by $\varphi \in H_0^1(\Omega)$ and $\psi \in \hat{H}_0^1(\Omega)$, respectively, results in the following weak formulation

$$-\sqrt{\alpha}\varepsilon(\nabla w, \nabla \varphi) + \alpha(\boldsymbol{\beta} \cdot \nabla y, \boldsymbol{\beta} \cdot \nabla \varphi) + (y, \varphi) = (F_2, \varphi), \tag{2.48}$$

$$(w, \psi) + \sqrt{\alpha}\varepsilon(\nabla y, \nabla \psi) - \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}} w, \psi >_{\Gamma_2} = \sqrt{\alpha}\varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}} f, \psi >_{\Gamma_2} . \tag{2.49}$$

Let $V_h \subset H^1(\Omega)$ be a finite element space consisting of piecewise linear polynomials. Set $V_h^0 = V_h \cap H_0^1(\Omega)$ and $\hat{V}_h^0 = V_h \cap \hat{H}_0^1(\Omega)$. Therefore the mixed finite element approximation of (2.42)-(2.43) is to find $(y_h, w_h) \in V_h^0 \times \hat{V}_h^0$ such that

$$-\sqrt{\alpha}\varepsilon(\nabla w_h, \nabla \varphi_h) + \alpha(\boldsymbol{\beta} \cdot \nabla y_h, \boldsymbol{\beta} \cdot \nabla \varphi_h) + (y_h, \varphi_h) = (F_2, \varphi_h), \tag{2.50}$$

$$(w_h,\psi_h)+\sqrt{\alpha}\varepsilon(\nabla y_h,\nabla\psi_h)-\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}w_h,\psi_h>_{\Gamma_2} \;=\; \sqrt{\alpha}\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}f,\psi_h>_{\Gamma_2}. \quad (2.51)$$

In the following, let us consider the error estimates of $y-y_h$ and $w-w_h$. Testing (2.48) with φ_h, (2.49) with ψ_h and subtracting (2.50)-(2.51) from the resulting equations yields the following error equations:

$$\begin{aligned}
-\sqrt{\alpha}\varepsilon(\nabla(w-w_h),\nabla\varphi_h)+\alpha(\boldsymbol{\beta}\cdot\nabla(y-y_h),\boldsymbol{\beta}\cdot\nabla\varphi_h)+(y-y_h,\varphi_h) &= 0, \quad (2.52)\\
(w-w_h,\psi_h)+\sqrt{\alpha}\varepsilon(\nabla(y-y_h),\nabla\psi_h)-\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}(w-w_h),\psi_h>_{\Gamma_2} &= 0. \quad (2.53)
\end{aligned}$$

Let y_I and w_I denote the Lagrange interpolation of y and w, respectively. We rewrite the error $y-y_h$ and $w-w_h$ as

$$\begin{aligned}
y-y_h &= y-y_I-(y_h-y_I)=\rho_y-\theta_y,\\
w-w_h &= w-w_I-(w_h-w_I)=\rho_w-\theta_w.
\end{aligned}$$

Then the error equation (2.52)-(2.53) can be rewritten as

$$\begin{aligned}
-\sqrt{\alpha}\varepsilon(\nabla\theta_w,\nabla\varphi_h)+\alpha(\boldsymbol{\beta}\cdot\nabla\theta_y,\boldsymbol{\beta}\cdot\nabla\varphi_h)+(\theta_y,\varphi_h) &= -\sqrt{\alpha}\varepsilon(\nabla\rho_w,\nabla\varphi_h)+\alpha(\boldsymbol{\beta}\cdot\nabla\rho_y,\boldsymbol{\beta}\cdot\nabla\varphi_h)\\
&+ (\rho_y,\varphi_h), \quad (2.54)\\
(\theta_w,\psi_h)+\sqrt{\alpha}\varepsilon(\nabla\theta_y,\nabla\psi_h)-\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\theta_w,\psi_h>_{\Gamma_2} &= (\rho_w,\psi_h)+\sqrt{\alpha}\varepsilon(\nabla\rho_y,\nabla\psi_h)\\
&- \varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\rho_w,\psi_h>_{\Gamma_2}. \quad (2.55)
\end{aligned}$$

Choosing $\varphi_h=\theta_y$ and $\psi_h=\theta_w$ in (2.54) and (2.55), respectively, and adding the resulting equations yields

$$\begin{aligned}
& \|\theta_y\|^2+\|\theta_w\|^2+\alpha\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2\\
= & -\sqrt{\alpha}\varepsilon(\nabla\rho_w,\nabla\theta_y)+\alpha(\boldsymbol{\beta}\cdot\nabla\rho_y,\boldsymbol{\beta}\cdot\nabla\theta_y)+(\rho_y,\theta_y),\\
+ & (\rho_w,\theta_w)+\sqrt{\alpha}\varepsilon(\nabla\rho_y,\nabla\theta_w)+\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\theta_w,\theta_w>_{\Gamma_2}-\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\rho_w,\theta_w>_{\Gamma_2},\\
= & \sum_{i=1}^{7}T_i. \quad (2.56)
\end{aligned}$$

Using Cauchy-Schwartz inequality together with an inverse inequality we obtain

$$\begin{aligned}
T_2 &\leqslant \frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\rho_y\|^2+\frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2,\\
T_3 &\leqslant \frac{1}{2}\|\rho_y\|^2+\frac{1}{2}\|\theta_y\|^2,\\
T_4 &\leqslant \frac{1}{2}\|\rho_w\|^2+\frac{1}{2}\|\theta_w\|^2,\\
T_6 &\leqslant C\frac{\varepsilon\underline{h}^{-1}}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\theta_w\|^2,\\
T_7 &\leqslant C\frac{\varepsilon\underline{h}^{-1}}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\theta_w\|^2+C\frac{\varepsilon}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\rho_w\|_{0,\Gamma_2}^2.
\end{aligned}$$

On general meshes a further application of an inverse inequality and the Cauchy-Schwartz inequality gives

$$T_1 \leqslant C\alpha\varepsilon^2\underline{h}^{-2}\|\nabla\rho_w\|^2+\delta\|\theta_y\|^2,$$

$$T_5 \leqslant C\alpha\varepsilon^2\underline{h}^{-2}\|\nabla\rho_y\|^2 + \delta\|\theta_w\|^2.$$

Inserting the estimates of $T_1 \sim T_7$ into (2.56) leads to

$$\begin{aligned} & \frac{1}{2}\|\theta_y\|^2 + \frac{1}{2}\|\theta_w\|^2 + \frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2 \\ \leqslant & \frac{1}{2}\|\rho_y\|^2 + \frac{1}{2}\|\rho_w\|^2 + C\frac{\varepsilon\underline{h}^{-1}}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\theta_w\|^2 \\ + & C\frac{\varepsilon}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\rho_w\|_{0,\Gamma_2}^2 + \frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\rho_y\|^2 \\ + & C\alpha\varepsilon^2\underline{h}^{-2}\|\nabla\rho_w\|^2 + C\alpha\varepsilon^2\underline{h}^{-2}\|\nabla\rho_y\|^2. \end{aligned}$$

Suppose that $\varepsilon < \gamma\underline{h}$ for a sufficiently small constant γ. Then we have

$$\frac{1}{2}\|\theta_y\|^2 + \delta\|\theta_w\|^2 + \frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2 \leqslant Ch^2(\|y\|_2^2 + \|w\|_2^2). \tag{2.57}$$

Using (2.57) and the estimates of Lagrange interpolation we arrive at

Theorem 2.3. *Let (y, w) and (y_h, w_h) be the solutions of (2.48)-(2.49) and (2.50)-(2.51), respectively. Assume that $\varepsilon < \gamma\underline{h}$. Then, there exists a constant C independent of h and ε such that*

$$\begin{aligned} \|y - y_h\| + \|w - w_h\| &\leqslant Ch(\|y\|_2 + \|w\|_2), \\ \|\boldsymbol{\beta}\cdot\nabla(y - y_h)\| &\leqslant C\frac{h}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + Ch\|y\|_2. \end{aligned}$$

According to [10] ifthe mesh $\mathcal{T}_h$ is rectangular, we have

$$\int_\Omega \nabla(\psi - \psi_I)\nabla v_h dx \leqslant Ch^2|\psi|_3|v_h|_1, \forall\psi \in H^3(\Omega), v_h \in V_h. \tag{2.58}$$

Utilizing (2.58) we then deduce

$$\begin{aligned} T_1 &\leqslant C\sqrt{\alpha}\varepsilon h^2\|w\|_3\|\theta_y\|_1 \\ &\leqslant C\sqrt{\alpha}\varepsilon h\|w\|_3\|\theta_y\|, \end{aligned} \tag{2.59}$$

and

$$\begin{aligned} T_5 &\leqslant C\sqrt{\alpha}\varepsilon h^2\|y\|_3\|\theta_w\|_1 \\ &\leqslant C\sqrt{\alpha}\varepsilon h\|y\|_3\|\theta_w\|. \end{aligned} \tag{2.60}$$

Moreover, if the mesh is uniform rectangular([10]), it can be proven that

$$\int_\Omega \sum_{i=1}^{2} \alpha_{ij}\partial_i(\psi - \psi_I)\partial_j v_h dx \leqslant Ch^2|\psi|_4\|v_h\|, \forall\psi \in H^4(\Omega), v_h \in V_h. \tag{2.61}$$

Then the term T_2 can be bounded as

$$T_2 \leqslant C\alpha h^2\|y\|_4\|\theta_y\|. \tag{2.62}$$

Therefore, it follows form (2.59), (2.62) and (2.60) together with the estimates of T_3, T_4, T_6, T_7 that

$$\frac{1}{2}\|\theta_y\|^2 + \frac{1}{2}\|\theta_w\|^2 + \frac{\alpha}{2}\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2 \leqslant \frac{1}{2}\|\rho_y\|^2 + \frac{1}{2}\|\rho_w\|^2 + C\frac{\varepsilon\underline{h}^{-1}}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\theta_w\|^2$$

$$
\begin{aligned}
&+ \quad C\frac{\varepsilon}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\rho_w\|^2_{0,\Gamma_2} + C\alpha^2h^4\|y\|_4^2 \\
&+ \quad C\alpha\varepsilon^2h^2\|w\|_3^2 + C\alpha\varepsilon^2h^2\|y\|_3^2 + \delta(\|\theta_y\|^2 + \|\theta_w\|^2).
\end{aligned}
$$

Again, suppose that $\varepsilon < \gamma\underline{h}$ for a sufficiently small constant γ. Then we derive

$$
\|\theta_y\|^2 + \|\theta_w\|^2 + \alpha\|\boldsymbol{\beta}\cdot\nabla\theta_y\|^2 \leqslant Ch^4(\|y\|_2^2 + \|w\|_2^2) + Ch^4(\|y\|_4^2 + \|w\|_3^2 + \|y\|_3^2). \quad (2.63)
$$

By (2.63) we have the following theorem.

Theorem 2.4. *Let (y, w) and (y_h, w_h) be the solutions of (2.48)-(2.49) and (2.50)-(2.51), respectively. Assume that $\varepsilon < \gamma\underline{h}$ and that the mesh is uniform rectangular. Then there exists a constant C independent of h and ε such that*

$$
\begin{aligned}
\|y - y_h\| + \|w - w_h\| &\leqslant Ch^2(\|y\|_4 + \|w\|_3), \\
\|\boldsymbol{\beta}\cdot\nabla(y - y_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|y\|_4 + \|w\|_3) + Ch\|y\|_2.
\end{aligned}
$$

Similar to one-dimensional problem, we derive the following governing equation for the adjoint state variable p in the two dimensional case:

$$
-\alpha\varepsilon^2\Delta^2 p + \alpha\boldsymbol{\beta}\cdot\nabla(\boldsymbol{\beta}\cdot\nabla p) - p = -\alpha\varepsilon\Delta z + \alpha\boldsymbol{\beta}\cdot\nabla z - \alpha f \triangleq -G_2. \quad (2.64)
$$

Defining $\vartheta = \sqrt{\alpha}\varepsilon\Delta p$ leads to

$$
\begin{aligned}
\vartheta - \sqrt{\alpha}\varepsilon\Delta p &= 0, && (2.65) \\
\sqrt{\alpha}\varepsilon\Delta\vartheta - \alpha\boldsymbol{\beta}\cdot\nabla(\boldsymbol{\beta}\cdot\nabla p) + p &= G_2. && (2.66)
\end{aligned}
$$

On the boundary we have

$$
p = 0, \quad \text{on} \quad \partial\Omega, \quad \vartheta = \sqrt{\alpha}z, \quad \text{on} \quad \Gamma_1, \quad (2.67)
$$

and

$$
\frac{\partial p}{\partial \mathbf{n}} = \frac{1}{(\boldsymbol{\beta}\cdot\mathbf{n})}(z - \frac{\vartheta}{\sqrt{\alpha}}), \quad \text{on} \quad \Gamma_2, \quad (2.68)
$$

where Γ_1 and Γ_2 are defined before.

Multiplying (2.65) and (2.66) by $\varphi \in H_0^1(\Omega)$ and $\psi \in \hat{H}_0^1(\Omega)$, respectively, yields

$$
\begin{aligned}
-\sqrt{\alpha}\varepsilon(\nabla\vartheta, \nabla\varphi) + \alpha(\boldsymbol{\beta}\cdot\nabla p, \boldsymbol{\beta}\cdot\nabla\varphi) + (p, \varphi) &= (G_2, \varphi), && (2.69) \\
(\vartheta, \psi) + \sqrt{\alpha}\varepsilon(\nabla p, \nabla\psi) + \varepsilon < \frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\vartheta, \psi >_{\Gamma_2} &= \sqrt{\alpha}\varepsilon < \frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}z, \psi >_{\Gamma_2}. && (2.70)
\end{aligned}
$$

Using the variational form (2.69)-(2.70) we define the following discrete scheme for the adjoint state variable p

$$
\begin{aligned}
-\sqrt{\alpha}\varepsilon(\nabla\vartheta_h, \nabla\varphi_h) + \alpha(\boldsymbol{\beta}\cdot\nabla p_h, \boldsymbol{\beta}\cdot\nabla\varphi_h) + (p_h, \varphi_h) &= (G_2, \varphi_h), && (2.71) \\
(\vartheta_h, \psi_h) + \sqrt{\alpha}\varepsilon(\nabla p_h, \nabla\psi_h) + \varepsilon < \frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\vartheta_h, \psi_h >_{\Gamma_2} &= \sqrt{\alpha}\varepsilon < \frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}z, \psi_h >_{\Gamma_2}. && (2.72)
\end{aligned}
$$

By arguments similar to that used in Theorem 2.4 we can deduce the following theorem results for the adjoint state variable p

Theorem 2.5. *Let (p, ϑ) and (p_h, ϑ_h) be the solutions of (2.69)-(2.70) and (2.71)-(2.72), respectively. Assume that $\varepsilon < \gamma\underline{h}$. Then, there exists a constant C independent of h and ε such that*

$$
\begin{aligned}
\|p - p_h\| + \|\vartheta - \vartheta_h\| &\leqslant Ch(\|p\|_2 + \|\vartheta\|_2), \\
\|\boldsymbol{\beta}\cdot\nabla(p - p_h)\| &\leqslant C\frac{h}{\sqrt{\alpha}}(\|p\|_2 + \|\vartheta\|_2) + Ch\|p\|_2.
\end{aligned}
$$

Theorem 2.6. *Let (p,ϑ) and (p_h,ϑ_h) be the solutions of (2.69)-(2.70) and (2.71)-(2.72), respectively. Assume that $\varepsilon<\gamma\underline{h}$ and that the mesh is uniform rectangular. Then there exists a constant C independent of h and ε such that*

$$\begin{aligned}\|p-p_h\|+\|\vartheta-\vartheta_h\| &\leqslant Ch^2(\|p\|_4+\|\vartheta\|_3),\\ \|\boldsymbol{\beta}\cdot\nabla(p-p_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_4+\|\vartheta\|_3)+Ch\|p\|_2.\end{aligned}$$

2.3 The discretization of control u

In this section, we consider the numerical approximation and corresponding error estimate for control variable u.

According to the first order optimality conditions we have

$$u=-\frac{p}{\alpha}.$$

Then we set

$$u_h=-\frac{p_h}{\alpha}.$$

Therefore the error of $u-u_h$ immediately follows from the estimates for $p-p_h$.

Theorem 2.7. *For one-dimensional problem suppose that $\varepsilon<\frac{h}{8\sqrt{3}}$. Then, there exists a constant C independent of h and ε such that*

$$\|u-u_h\| \leqslant C(\|p\|_2+\|\vartheta\|_2)\frac{h^2}{\alpha}.$$

Theorem 2.8. *For two-dimensional problems assume that $\varepsilon<\gamma\underline{h}$ for a sufficient small positive constant γ. Then there exists a constant C independent of h and ε such that*

$$\|u-u_h\|\leqslant\begin{cases} C(\|p\|_2+\|\vartheta\|_2)\frac{h}{\alpha}, & \text{on a general mesh},\\ C(\|p\|_4+\|\vartheta\|_3)\frac{h^2}{\alpha}, & \text{on a uniform rectangular mesh.}\end{cases}$$

Moreover, we consider another different discretization for control variable u.

For the one dimensional case we get by (2.4)

$$\begin{aligned}u &= -\varepsilon y''+y'-f\\ &= -\frac{w}{\sqrt{\alpha}}+y'-f.\end{aligned} \tag{2.73}$$

We approximate the control variable u as follows

$$u_h = -\frac{w_h}{\sqrt{\alpha}}+y_h'-f_I, \tag{2.74}$$

where f_I is the Lagrange interpolation of f.

Then by Theorem 2.1 we can deduce the following error estimate for $u-u_h$

$$\begin{aligned}\|u-u_h\| &\leqslant \|\frac{w-w_h}{\sqrt{\alpha}}\|+\|y'-y_h'\|+\|f-f_I\|\\ &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|y\|_2+\|w\|_2)+Ch\|y\|_2+Ch\|f\|_1,\end{aligned}$$

which can be stated in the following theorem.

Theorem 2.9. *Let (y, w) and (y_h, w_h) be the solutions of (2.15)-(2.16) and (2.17)-(2.18), respectively. Assume that $\varepsilon < \frac{h}{8\sqrt{3}}$. Then, there exists a constant C independent of h and ε such that*

$$\|u - u_h\| \leqslant C\frac{h^2}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + Ch\|y\|_2 + Ch\|f\|_1.$$

Similarly to one dimensional problem we derive by (2.37)

$$\begin{aligned} u &= -\varepsilon\Delta y + \boldsymbol{\beta}\cdot\nabla y - f \\ &= -\frac{w}{\sqrt{\alpha}} + \boldsymbol{\beta}\cdot\nabla y - f. \end{aligned} \tag{2.75}$$

The control variable u can be approximated as follows

$$u_h = -\frac{w_h}{\sqrt{\alpha}} + \boldsymbol{\beta}\cdot\nabla y_h - f_I. \tag{2.76}$$

Then we have the following error estimate for $u - u_h$

$$\begin{aligned} \|u - u_h\| &\leqslant \|\frac{w - w_h}{\sqrt{\alpha}}\| + \|\boldsymbol{\beta}\cdot(\nabla y - \nabla y_h)\| + \|f - f_I\| \\ &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + C\frac{h}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + Ch\|f\|_1. \end{aligned}$$

We thus have the following

Theorem 2.10. *Let (y, w) and (y_h, w_h) be the solutions of (2.48)-(2.49) and (2.50)-(2.51), respectively. Assume that $\varepsilon < \gamma\underline{h}$. Then, there exists a constant C independent of h and ε such that*

$$\|u - u_h\| \leqslant C\frac{h}{\sqrt{\alpha}}(\|y\|_2 + \|w\|_2) + Ch\|y\|_2 + Ch\|f\|_1.$$

3 Optimal control problem with constraints

In this section we only consider the following two dimensional convection diffusion optimal control problems with control constraints;

$$\min_{u\in U_{ad}} J(y, u) = \frac{1}{2}\int_\Omega (y - z)^2 dx + \frac{\alpha}{2}\int_\Omega u^2 dx \tag{3.1}$$

subject to

$$-\varepsilon\Delta y + \boldsymbol{\beta}\cdot\nabla y + \nu y = f + u, \quad \text{in } \Omega, \tag{3.2}$$

$$y = 0, \quad \text{on } \partial\Omega. \tag{3.3}$$

Here $\Omega \subset R^2$ is a bounded domain, and α is a positive constant. f and z are given functions. $U_{ad} = \{v \in L^2(\Omega); v \geqslant 0\}$. $\boldsymbol{\beta}$ is constant vector, $\nu > 0$ denotes the reaction coefficient and $\varepsilon > 0$ denotes the diffusion coefficient.

By standard argument we derive the following first order optimality conditions for optimal control problems (3.1)-(3.3);

$$-\varepsilon\Delta y + \boldsymbol{\beta}\cdot\nabla y + \nu y = f + u, \quad x \in \Omega, \tag{3.4}$$

$$-\varepsilon\Delta p - \boldsymbol{\beta}\cdot\nabla p + \nu p = y - z, \quad x \in \Omega, \tag{3.5}$$

$$u - \max\{0, -\frac{p}{\alpha}\} \quad = 0, \qquad x \in \Omega, \tag{3.6}$$
$$y = 0,\ p \quad = 0, \qquad x \in \partial\Omega. \tag{3.7}$$

By (3.5) we have

$$y = z - \varepsilon\Delta p - \boldsymbol{\beta} \cdot \nabla p + \nu p.$$

Inserting the above equality into the state equation (3.5) yields

$$\varepsilon^2\Delta^2 p - \boldsymbol{\beta} \cdot \nabla(\boldsymbol{\beta} \cdot \nabla p) - 2\nu\varepsilon\Delta p + \nu^2 p - \max\{0, -\frac{p}{\alpha}\} = f - \boldsymbol{\beta} \cdot \nabla z - \nu z + \varepsilon\Delta z \triangleq M. \tag{3.8}$$

By introducing $\mathscr{W} = \varepsilon\Delta p$ we can decompose (3.8) into the following systems

$$\mathscr{W} - \varepsilon\Delta p \quad = \quad 0, \tag{3.9}$$
$$\varepsilon\Delta\mathscr{W} - \boldsymbol{\beta} \cdot \nabla(\boldsymbol{\beta} \cdot \nabla p) - 2\nu\varepsilon\Delta p + \nu^2 p - \max\{0, -\frac{p}{\alpha}\} \quad = \quad M. \tag{3.10}$$

For the boundary condition we have

$$p = 0, \quad \boldsymbol{\beta} \cdot \nabla p = z - \mathscr{W}, \quad \text{on} \ \ \partial\Omega.$$

Further we have

$$p = 0, \quad \text{on} \ \ \partial\Omega, \quad \mathscr{W} = z, \quad \text{on} \ \ \Gamma_1, \tag{3.11}$$

and

$$\frac{\partial p}{\partial \mathbf{n}} = \frac{1}{(\boldsymbol{\beta} \cdot \mathbf{n})}(z - \mathscr{W}), \quad \text{on} \ \ \Gamma_2, \tag{3.12}$$

where Γ_1 and Γ_2 are defined before.

Multiplying (3.9) and (3.10) by $\varphi \in H_0^1(\Omega)$ and $\psi \in \hat{H}_0^1(\Omega)$, respectively, yields

$$-\varepsilon(\nabla\mathscr{W}, \nabla\varphi) + (\boldsymbol{\beta} \cdot \nabla p, \boldsymbol{\beta} \cdot \nabla\varphi) + 2\varepsilon(\nu\nabla p, \nabla\varphi)$$
$$+(\nu^2 p, \varphi) - (\max\{0, -\frac{p}{\alpha}\}, \varphi) = (M, \varphi), \tag{3.13}$$
$$(\mathscr{W}, \psi) + \varepsilon(\nabla p, \nabla\psi) + \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}}\mathscr{W}, \psi >_{\Gamma_2} = \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}} z, \psi >_{\Gamma_2}. \tag{3.14}$$

The following discrete scheme for adjoint state variable p can be defined by using the variational form (3.13)-(3.14)

$$-\varepsilon(\nabla\mathscr{W}_h, \nabla\varphi_h) + (\boldsymbol{\beta} \cdot \nabla p_h, \boldsymbol{\beta} \cdot \nabla\varphi_h) + 2\varepsilon(\nu\nabla p_h, \nabla\varphi_h)$$
$$+(\nu^2 p_h, \varphi_h) - (\max\{0, -\frac{p_h}{\alpha}\}, \varphi_h) = (M, \varphi_h), \tag{3.15}$$
$$(\mathscr{W}_h, \psi_h) + \varepsilon(\nabla p_h, \nabla\psi_h) + \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}}\mathscr{W}_h, \psi_h >_{\Gamma_2} = \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}} z, \psi_h >_{\Gamma_2}. \tag{3.16}$$

In order to derive the error estimates according to [11] we introduce the following auxiliary problems

$$-\varepsilon(\nabla\widetilde{\mathscr{W}}_h, \nabla\varphi_h) + (\boldsymbol{\beta} \cdot \nabla\tilde{p}_h, \boldsymbol{\beta} \cdot \nabla\varphi_h) + 2\varepsilon(\nu\nabla\tilde{p}_h, \nabla\varphi_h)$$
$$+(\nu^2\tilde{p}_h, \varphi_h) - (\max\{0, -\frac{p}{\alpha}\}, \varphi_h) = (M, \varphi_h), \tag{3.17}$$
$$(\widetilde{\mathscr{W}}_h, \psi_h) + \varepsilon(\nabla\tilde{p}_h, \nabla\psi_h) + \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}}\widetilde{\mathscr{W}}_h, \psi_h >_{\Gamma_2} = \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}} z, \psi_h >_{\Gamma_2}. \tag{3.18}$$

By similar arguments as used in the unconstrained case we deduce the following error estimates.

Theorem 3.1. *Let $(p, \mathscr{W})$ and $(\tilde{p}_h, \widetilde{\mathscr{W}_h})$ be the solutions of (3.13)-(3.14) and (3.17)-(3.18), respectively. Assume that $\varepsilon < \gamma \underline{h}$. Then, there exists a constant C independent of h and ε such that*

$$\begin{aligned}\|p - \tilde{p}_h\| + \|\mathscr{W} - \widetilde{\mathscr{W}_h}\| &\leqslant Ch(\|p\|_2 + \|\mathscr{W}\|_2), \\ \|\boldsymbol{\beta} \cdot \nabla(p - \tilde{p}_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_2 + \|\mathscr{W}\|_2) + Ch\|p\|_2.\end{aligned}$$

Theorem 3.2. *Let $(p, \mathscr{W})$ and $(\tilde{p}_h, \widetilde{\mathscr{W}_h})$ be the solutions of (3.13)-(3.14) and (3.17)-(3.18), respectively. Assume that $\varepsilon < \gamma \underline{h}$ and that the mesh is uniform rectangular. Then there exists a constant C independent of h and ε such that*

$$\begin{aligned}\|p - \tilde{p}_h\| + \|\mathscr{W} - \widetilde{\mathscr{W}_h}\| &\leqslant Ch^2(\|p\|_4 + \|\mathscr{W}\|_3), \\ \|\boldsymbol{\beta} \cdot \nabla(p - \tilde{p}_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_4 + \|\mathscr{W}\|_3) + Ch\|p\|_2.\end{aligned}$$

Utilizing (3.15)-(3.16) and (3.17)-(3.18) we have

$$\begin{aligned}&-\varepsilon(\nabla(\mathscr{W}_h - \widetilde{\mathscr{W}_h}), \nabla\varphi_h) + (\boldsymbol{\beta} \cdot \nabla(p_h - \tilde{p}_h), \boldsymbol{\beta} \cdot \nabla\varphi_h) + 2\varepsilon(\nu\nabla(p_h - \tilde{p}_h), \nabla\varphi_h) \\ &\quad +(\nu^2(p_h - \tilde{p}_h), \varphi_h) - (\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \varphi_h) = 0, \qquad (3.19)\\ &((\mathscr{W}_h - \widetilde{\mathscr{W}_h}), \psi_h) + \varepsilon(\nabla(p_h - \tilde{p}_h), \nabla\psi_h) + \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}}(\mathscr{W}_h - \widetilde{\mathscr{W}_h}), \psi_h >_{\Gamma_2} = 0. \qquad (3.20)\end{aligned}$$

Let $\zeta = \mathscr{W}_h - \widetilde{\mathscr{W}_h}$ and $\chi = p_h - \tilde{p}_h$. Testing (3.19) with $\varphi_h = \zeta$, (3.20) with $\psi_h = \chi$ and adding the resulting equations together leads to

$$\begin{aligned}&\|\zeta\|^2 + \|\boldsymbol{\beta} \cdot \nabla\chi\|^2 + 2\varepsilon\nu\|\nabla\chi\|^2 + \|\nu\chi\|^2 \\ &= (\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \chi) - \varepsilon < \frac{1}{\boldsymbol{\beta} \cdot \mathbf{n}}\zeta, \zeta >_{\Gamma_2}. \qquad (3.21)\end{aligned}$$

Note that

$$\begin{aligned}&(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, p_h - \tilde{p}_h) \\ =\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \frac{1}{\alpha}p_h - \frac{1}{\alpha}\tilde{p}_h) \\ =\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \frac{1}{\alpha}p_h - \frac{1}{\alpha}p + \frac{1}{\alpha}p - \frac{1}{\alpha}\tilde{p}_h) \\ =\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \frac{1}{\alpha}p_h - \frac{1}{\alpha}p) + (\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, p - \tilde{p}_h) \\ =\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \frac{1}{\alpha}p_h + \max\{0, -\frac{p_h}{\alpha}\}) \\ -\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}) \qquad (3.22)\\ +\ & \alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, -\max\{0, -\frac{p}{\alpha}\} - \frac{1}{\alpha}p) \\ +\ & (\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, p - \tilde{p}_h).\end{aligned}$$

It is easy to prove that

$$\alpha(\max\{0, -\frac{p_h}{\alpha}\} - \max\{0, -\frac{p}{\alpha}\}, \frac{1}{\alpha}p_h + \max\{0, -\frac{p_h}{\alpha}\})$$

$$
\begin{aligned}
&+ \quad \alpha(\max\{0,-\frac{p_h}{\alpha}\}-\max\{0,-\frac{p}{\alpha}\},-\max\{0,-\frac{p}{\alpha}\}-\frac{1}{\alpha}p) \\
&= \quad \begin{cases} (0,p_h)+(0,-p), & p\geqslant 0,\ p_h\geqslant 0, \\ (p_h,0)+\frac{1}{\alpha}(p_h,p), & p\geqslant 0,\ p_h<0, \\ \frac{1}{\alpha}(p_h,p)+\frac{1}{\alpha}(p,0), & p<0,\ p_h\geqslant 0, \\ 0, & p<0,\ p_h<0, \end{cases} \\
&\leqslant \quad 0. \qquad (3.23)
\end{aligned}
$$

It follows from (3.21), (3.22) and (3.23) that

$$
\begin{aligned}
&\|\zeta\|^2+\|\boldsymbol{\beta}\cdot\nabla\chi\|^2+2\varepsilon\nu\|\nabla\chi\|^2+\|\nu\chi\|^2+\alpha\|\max\{0,-\frac{p_h}{\alpha}\}-\max\{0,-\frac{p}{\alpha}\}\|^2 \\
&= (\max\{0,-\frac{p_h}{\alpha}\}-\max\{0,-\frac{p}{\alpha}\},p-\tilde{p}_h)-\varepsilon<\frac{1}{\boldsymbol{\beta}\cdot\mathbf{n}}\zeta,\zeta>_{\Gamma_2} \\
&\leqslant C\|p-\tilde{p}_h\|^2+\frac{\alpha}{2}\|\max\{0,-\frac{p_h}{\alpha}\}-\max\{0,-\frac{p}{\alpha}\}\|^2+C\frac{\varepsilon h^{-1}}{\min(|\boldsymbol{\beta}\cdot\mathbf{n}|)}\|\zeta\|^2 \qquad (3.24)
\end{aligned}
$$

Suppose that $\varepsilon<\gamma h$ for a sufficiently small constant γ. Then we deduce

$$
\begin{aligned}
&\|\zeta\|^2+\|\boldsymbol{\beta}\cdot\nabla\chi\|^2+2\varepsilon\nu\|\nabla\chi\|^2+\|\nu\chi\|^2+\alpha\|\max\{0,-\frac{p_h}{\alpha}\}-\max\{0,-\frac{p}{\alpha}\}\|^2 \\
&\leqslant C\|p-\tilde{p}_h\|^2. \qquad (3.25)
\end{aligned}
$$

Combining (3.25) with Theorem 3.1 and 3.2 we arrive at

Theorem 3.3. *Let $(p,\mathscr{W})$ and $(p_h,\mathscr{W}_h)$ be the solutions of (3.13)-(3.14) and (3.15)-(3.16), respectively. Assume that $\varepsilon<\gamma\underline{h}$. Then, there exists a constant C independent of h and ε such that*

$$
\begin{aligned}
\|p-p_h\|+\|\mathscr{W}-\mathscr{W}_h\| &\leqslant Ch(\|p\|_2+\|\mathscr{W}\|_2), \\
\|\boldsymbol{\beta}\cdot\nabla(p-p_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_2+\|\mathscr{W}\|_2)+Ch\|p\|_2.
\end{aligned}
$$

Theorem 3.4. *Let $(p,\mathscr{W})$ and $(p_h,\mathscr{W}_h)$ be the solutions of (3.13)-(3.14) and (3.15)-(3.16), respectively. Assume that $\varepsilon<\gamma\underline{h}$ and that the mesh is uniform rectangular. Then there exists a constant C independent of h and ε such that*

$$
\begin{aligned}
\|p-p_h\|+\|\mathscr{W}-\mathscr{W}_h\| &\leqslant Ch^2(\|p\|_4+\|\mathscr{W}\|_3), \\
\|\boldsymbol{\beta}\cdot\nabla(p-p_h)\| &\leqslant C\frac{h^2}{\sqrt{\alpha}}(\|p\|_4+\|\mathscr{W}\|_3)+Ch\|p\|_2.
\end{aligned}
$$

Using the first order optimality conditions we approximate the control variable according to [8] as

$$
u_h=\max\big\{0,-\frac{p_h}{\alpha}\big\}. \qquad (3.26)
$$

Then it is easy to see that $u_h\in U_{ad}$. In general, u_h is not a finite element function corresponding to the mesh $\mathscr{T}_h$, especially on triangles containing the discrete free boundary. For the control variable we have the following error estimate.

Theorem 3.5. *Assume that $\varepsilon<\gamma\underline{h}$ for a sufficiently small positive constant γ. Then there exists a constant C independent of h and ε such that*

$$
\|u-u_h\|\leqslant\begin{cases} C(\|p\|_2+\|\mathscr{W}\|_2)\frac{h}{\alpha}, & \text{on a general mesh,} \\ \\ C(\|p\|_4+\|\mathscr{W}\|_3)\frac{h^2}{\alpha}, & \text{on a uniform rectangular mesh.} \end{cases}
$$

In constrained problem our method doesn't work for the state variable. To discretize the state variable we can adopt some stabilized finite element methods, such as the local projection stabilized method([1]) and the edge stabilization Galerkin method([4]).

To our best knowledge utilizing such methods to discretize the state variable only leads to $O(h^{\frac{3}{2}})$ errors with piecewise linear polynomials approximation.

Proceeding similar to [4] we apply the edge stabilization Galerkin method to discretize the state equation. To control the convective derivative of the discrete solution sufficiently we introduce a stabilization form ([13])

$$S(v_h, w_h) = \sum_{e \in \mathscr{E}_h} \int_e \varkappa h_e^2 [\nabla v_h \cdot \mathbf{n}][\nabla w_h \cdot \mathbf{n}] ds, \tag{3.27}$$

where $\mathscr{E}_h$ denotes the collection of interior edges of the triangles in T^h, h_e is the size of the edge e, $[q]$ denotes the jump of q across e for $e \in \mathscr{E}_h$

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

$\mathbf{n}$ is the outward unit normal. $\varkappa$ is a constant independent of h and ε.

Using above stabilization forms, an edge stabilization Galerkin approximation of the state equation (3.4) is as follows:

$$A(y_h, w_h) + S(y_h, w_h) = (f + u_h, w_h), \qquad \forall w_h \in V_h^0, \tag{3.28}$$

where $A(\cdot, \cdot)$ denotes the bilinear form given by

$$A(y, w) = (\varepsilon \nabla y, \nabla w) + (\boldsymbol{\beta} \cdot \nabla y, w) + (\nu y, w), \qquad y, w \in H_0^1(\Omega).$$

Theorem 3.6. *Let (y, u) and (y_h, u_h) be the solutions of the equations (3.4)-(3.6) and (3.28)-(3.26), respectively. Assume that all conditions of Theorem 3.5 are valid. Then we have the following estimates*

$$\| y - y_h \|_* \leqslant \begin{cases} C(h + h^{3/2} + h\varepsilon^{1/2}), & \text{on a general mesh,} \\ C(h^2 + h^{3/2} + h\varepsilon^{1/2}), & \text{on a uniform rectangular mesh.} \end{cases}$$

where

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

Proof. Let $\tilde{y}_h$ be the solution of the following equation

$$A(\tilde{y}_h, w_h) + S(\tilde{y}_h, w_h) = (f + u, w_h), \quad \forall w_h \in V_h^0.$$

We observe that $\tilde{y}_h$ is the edge stabilization Galerkin approximation of y, and by the stability property of $A(\cdot, \cdot) + S(\cdot, \cdot)$ (see,e.g.,[13]) we derive

$$\| y_h - \tilde{y}_h \|_* \leqslant C \| u - u_h \| . \tag{3.29}$$

Again proceeding similarly as in [13] leads to

$$\| y - \tilde{y}_h \|_* \leqslant C(h^{3/2} + h\varepsilon^{1/2}). \tag{3.30}$$

Then (3.29), (3.30) as well as the results of Theorem 3.5 imply that

$$\begin{aligned} & \| y - y_h \|_{*,\Omega} \\ \leqslant \; & C(h^{3/2} + h\varepsilon^{1/2}) + C \| u - u_h \| \\ \leqslant \; & \begin{cases} C(h + h^{3/2} + h\varepsilon^{1/2}), & \text{on a general mesh,} \\ C(h^2 + h^{3/2} + h\varepsilon^{1/2}), & \text{on a uniform rectangular mesh.} \end{cases} \end{aligned}$$

□

4 Numerical examples

The goal of this section is to carry out numerical examples to illustrate our analytical findings. Consider the following control problem

$$\min_{u\in U} J(y,u) = \frac{1}{2}\int_\Omega (y-z)^2dx + \frac{1}{2}\int_\Omega u^2dx \tag{4.1}$$

subject to

$$-\varepsilon\Delta y + \boldsymbol{\beta}\cdot\nabla y + \nu y = f+u, \quad \text{in } \Omega = [0,1]\times[0,1], \tag{4.2}$$
$$y = y_0, \quad \text{on } \partial\Omega. \tag{4.3}$$

Example 4.1. *Let $\varepsilon = 10^{-3}$, $\boldsymbol{\beta} = [1,1]^T$ and $\nu = 0$ in (4.2). The exact solutions are chosen as*

$$y = 16x_1(1-x_1)x_2(1-x_2)\times(\frac{1}{2}+\frac{1}{\pi}arctan((\frac{1}{8}-2(x_1-0.5)^2-2(x_2-0.5)^2)/\sqrt{\varepsilon})),$$
$$p = \sin(2\pi x_1)\sin(2\pi x_2).$$

The corresponding desired state z, and desired righthand side f can be calculated from the exact solutions and the governing equations.

Firstly we apply the mixed finite element method proposed in section 2 to solve numerically a one-dimensional problem. The corresponding errors for y, w and p, ϑ are presented in Table 4.1, which imply

$$\|y-y_h\| + \|p-p_h\| = O(h^2).$$

The figures for y and w are shown in Figure 4.1 and 4.2, respectively.

Table 4.1. Error of state y, w and adjoint state p , ϑ for the one-dimensional problem.

h	$\|y-y_h\|_{0,\Omega}$	order	$\|w-w_h\|_{0,\Omega}$	order	$\|p-p_h\|_{0,\Omega}$	order	$\|\vartheta-\vartheta_h\|$	order
$\frac{1}{60}$	0.0034	\	0.0086	\	3.0748e-4	\	1.3370e-5	\
$\frac{1}{120}$	6.6531e-4	2.3534	0.0026	1.7258	7.6810e-5	2.0011	3.3452e-6	1.9988
$\frac{1}{240}$	1.6406e-4	2.0198	7.1743e-4	1.8576	1.9199e-5	2.0003	8.3648e-7	1.9997
$\frac{1}{480}$	4.0879e-5	2.0048	1.8334e-4	1.9683	4.7994e-6	2.0001	2.0913e-7	1.9999

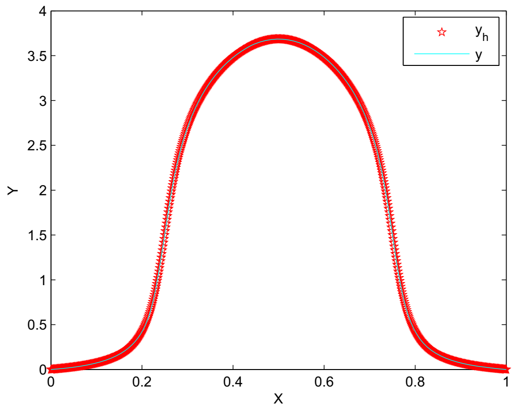

Figure 4.1. y together with y_h.

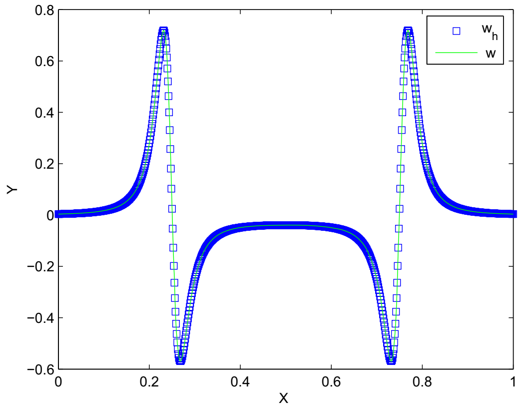

Figure 4.2. w together with w_h.

Secondly, we utilize the mixed finite element method proposed in section 2 to solve numerically a two-dimensional problem on uniform triangle mesh. Table 4.2 displays the errors of state y, w and adjoint state p, ϑ, from which we observe that

$$\|y - y_h\| + \|p - p_h\| = O(h^2).$$

The figures of y_h and w_h are presented in Figure 4.3 and 4.4, respectively.

Table 4.2. Error of state y, w and adjoint state p , ϑ for the two-dimensional problem on uniform triangle meshes.

h	$\|y-y_h\|_{0,\Omega}$	order	$\|w-w_h\|_{0,\Omega}$	order	$\|p-p_h\|_{0,\Omega}$	order	$\|\vartheta-\vartheta_h\|$	order
$\frac{1}{50}$	0.0019	\	0.0089	\	0.0011	\	1.2698e-004	\
$\frac{1}{60}$	0.0013	2.0814	0.0062	1.9828	8.0058e-004	1.7427	8.7934e-005	2.0154
$\frac{1}{70}$	9.5902e-004	1.9734	0.0043	2.3739	5.9147e-004	1.9638	6.4557e-005	2.0048
$\frac{1}{80}$	7.3740e-004	1.9679	0.0032	2.2127	4.5575e-004	1.9521	4.9559e-005	1.9799

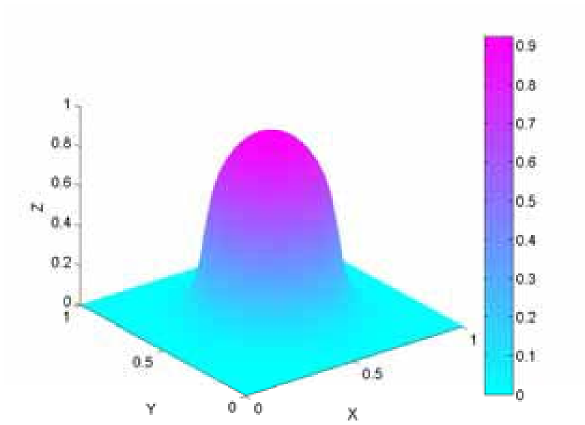

Figure 4.3. y_h.

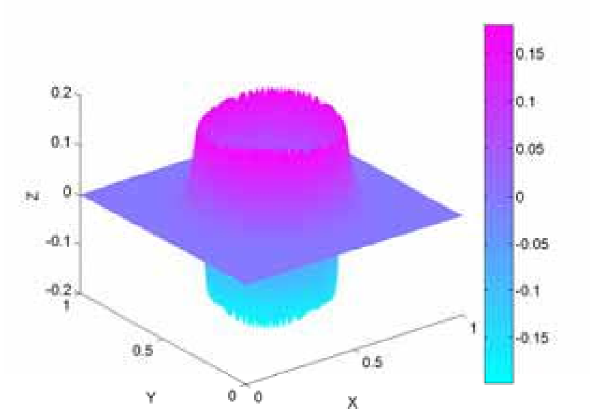

Figure 4.3. w_h.

Example 4.2. *Let $\alpha = 0.1$ in (4.1), $\varepsilon = 10^{-3}$, $\boldsymbol{\beta} = [1,1]^T$ and $\nu = 1$ in (4.2). The exact solutions are chosen as*

$$\begin{aligned} y &= \frac{2}{\pi}(atan(100(-0.5x_1 + x_2 - 0.25))), \\ p &= 16x_1(1-x_1)x_2(1-x_2) \times (\frac{1}{2} + \frac{1}{\pi}arctan((\frac{1}{8} - 2(x_1-0.5)^2 - 2(x_2-0.5)^2)/\sqrt{\varepsilon})), \\ u &= \max\big\{-5, \min\big\{-2.5, -\frac{p}{\alpha}\big\}\big\}. \end{aligned}$$

The corresponding desired state z, and desired righthand side f can be calculated from the exact solutions and the governing equations.

In this example we consider a control constrained problem. We use the mixed finite element method proposed in this paper to solve numerically the adjoint state variable, while the edge stabilization Galerkin method is applied to approximate the state variable. The errors for state y, adjoint state p and control u are displayed in Table 4.3, respectively. The figures of discrete state y and control u are shown in Figure 4.5 and 4.6.

Table 4.3. Error of state y, adjoint state p, ϑ and control u for the two-dimensional problem on uniform triangle meshes.

h	$\|y-y_h\|$	order	$\|p-p_h\|_{0,\Omega}$	order	$\|\vartheta-\vartheta_h\|_{0,\Omega}$	order	$\|u-u_h\|$	order
$\frac{1}{50}$	0.0115	\	0.0018	\	0.0088	\	0.0085	\
$\frac{1}{60}$	0.0075	2.3445	0.0013	1.7849	0.0062	1.9208	0.0064	1.5564
$\frac{1}{70}$	0.0052	2.3759	9.4974e-004	2.0365	0.0043	2.3739	0.0047	2.0028
$\frac{1}{80}$	0.0038	2.3489	7.2985e-004	1.9722	0.0032	2.2127	0.0035	2.2077

From Table 4.3 we observe that

$$\|u-u_h\| + \|p-p_h\| = O(h^2),$$

which are in agreement with our theoretical analysis in section 3. Moreover, we find that

$$\|y-y_h\| = O(h^2),$$

which is better than predicted by our theoretical analysis.

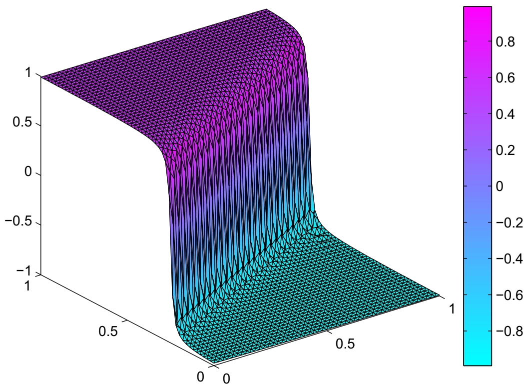

Figure 4.5. y_h.

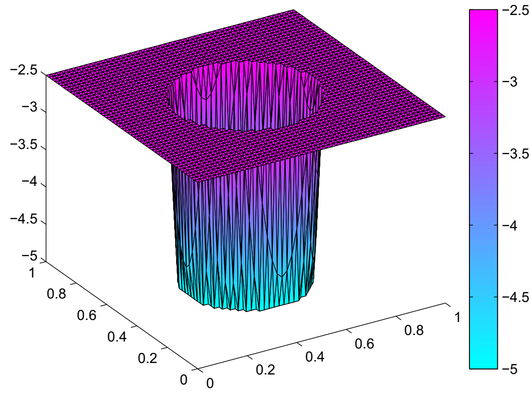

Figure 4.6. u_h.

In summary, from a numerical point of view we conclude that for unconstrained problems the mixed finite element method proposed in this paper has $O(h^2)$ convergence order for the state variable y, the adjoint state variable p as well as the control variable u, and for constrained problems this method has $O(h^2)$ convergence order for the adjoint state variable p and the control variable u.

Acknowledgments. The first author and the third author gratefully acknowledge the support of DFG SPP1253. The second author acknowledges the support of the National Basic Research Program under the Grant 2010CB731505.

References

[1] R. Becker and B.Vexler, Optimal control of the convection-diffusion equation using stabilized finite element methods, *Numer. Math.,* 106(3), 349–367, 2007.

[2] G. Lube and B. Tews, Optimal control of sigularly pertuberd advection diffusion reaction problems, *Mathematical Models and Methods in Applied Sciences*, DOI No: 10.1142/S0218202510004271.

[3] 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.

[4] 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, *J. Computational and Applied Mathematics,* 223, 198–217, 2009.

[5] N. Yan and Z. Zhou, A RT mixed FEM/DG scheme for optimal control governed by convection diffusion equations, *Journal of Scientific Computing,* 41, 273–299, 2009.

[6] N. Yan and Z. Zhou, A priori and a posteriori error estimates of streamline diffusion finite element method for optimal control problem governed by convection dominated diffusion equation, *Numerical Mathematics: Theory, Method and Applications,* 1(3), 297–320, 2008.

[7] J. Zhu and Q. Zeng, A mathematical theoretical frame for control of air pollution, *Sci. China, Ser. D*, 32, 864–870, 2002

[8] M. Hinze, A variational discretization concept in control constrained optimization: the linear quadratic case, Comput. Optim. Appl., 30, 45–61, 2005.

[9] A. Martínez, C. Rodríguez and M.E. Vázquez-Méndez, Theoretical and Numerical Analysis of an Optimal Control Problem Related to Wastewater Treatment, SIAM J. Control Optim., 38(5), 1534–1553, 2000.

[10] Q. Lin and N. Yan, Structure and analysis for efficient finite element methods. Baoding: Publishers of Hebei University(in Chinese), 1996.

[11] L. Chang, W. Gong and N. Yan, Finite element method for a nonsmooth elliptic equation. Front. Math. China 5(2), 191–209, 2010.

[12] M. Hinze, N. Yan and Z. Zhou, Variational discretization for optimal control governed by convection dominated diffusion equations. J. Comput. Math., 27, 237-253, 2009.

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