INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 7, Number 1, Pages 87–107

DISCONTINUOUS GALERKIN FINITE ELEMENT METHOD WITH INTERIOR PENALTIES FOR CONVECTION DIFFUSION OPTIMAL CONTROL PROBLEM

TONGJUN SUN

(Communicated by Wenbin Liu)

Abstract. In this paper, a discontinuous Galerkin finite element method with interior penalties for convection-diffusion optimal control problem is studied. A semi-discrete time DG scheme for this problem is presented. We analyze the stability of this scheme, and derive a priori and a posteriori error estimates for both the state and the control approximation.

Key Words. discontinuous Galerkin method, convection-diffusion equation, optimal control problem, a priori error estimates, a posteriori error estimates

1. Introduction

Finite element approximation of optimal control problems has been an important topic in engineering design work. There has been extensive theoretical and numerical studies for standard finite element approximation of various optimal control problems. For instance, for the optimal control problems governed by some linear elliptic or parabolic state equations, a priori error estimates of the finite element approximation were established long ago, see [1, 2, 3, 4, 5]. Furthermore, a priori error estimates were established for the finite element approximation of some important flow control problems in [6]. Some recent progress in a priori error estimates can be found in [7, 8] and in [9, 10, 11, 12], for a posteriori error estimates. Systematic introduction of the finite element method for PDEs and optimal control problems can be found in, for example, [13], [14] and [15].

In recent years, the discontinuous Galerkin methods have been proved very useful in solving a large range of computational fluid problems ([16, 17, 18]). They are preferred over standard continuous Galerkin methods because of their flexibility in approximating globally rough solutions, their local mass conservation, their possible definition on unstructured meshes, their potential for error control and mesh adaptation.

The idea of using penalty terms in a finite element method is not new. Baker [19] was the first one who used interior penalty with nonconforming elements for elliptic equations. Douglas and Dupont [20] analyzed a method which used interior penalties on the derivatives with conforming elements for linear elliptic and parabolic problems. Inspired by [19], Wheeler [21] presented an interior penalty method for second order linear elliptic equations. Closest to [21], Arnold [22] formulated a semi-discrete discontinuous Galerkin method with interior penalty for second order nonlinear parabolic equations.

Received by the editors September 9, 2008 and, in revised form, July 5, 2009.
2000 *Mathematics Subject Classification.* 65N30, 49J20 .
This research was supported by the NSF of China(No.10571108) and SRF for ROCS, SEM .

These methods [20, 21, 22] generalized a method by Nitsche [23] for treating Dirichlet boundary condition by the introduction of penalty terms on the boundary of the domain. Applications of these methods to flow in porous media were presented by Douglas, Wheeler, Darlow and Kendall in [24]. These methods frequently referred to as interior penalty Galerkin schemes.

In general, penalty terms are weighted L^2 inner products of the jumps in the function values across element edges. The primary motivation of including interior penalties is to impose approximate continuity. These terms enable closer approximation of solutions which varies in character from one element to another and allow the incorporation of partial knowledge of the solution into the scheme. Numerical experiments have clearly demonstrated the value of penalties for solving certain problems (see, e.g., [20]). New applications of discontinuous Galerkin method with interior penalties to nonlinear parabolic equations were introduced and analyzed by Rivière and Wheeler ([17, 25, 26]). It was shown that the method in ([17, 25, 26]) was elementwise conservative, and a priori and a posteriori error estimates in higher dimensions were derived.

Optimal control for convection-diffusion equation is widely met in practical applications. For example, in Environmental Sciences, some phenomena modelled by linear convection-diffusion partial differential equations are often studied to investigate the distribution forecast of pollutants in water or in atmosphere. In this context it might be of interest to regulate the source term of the convection-diffusion equation so that the solution is as near as possible to a desired one, e.g., to operate the emission rates of industrial plants to keep the concentration of pollutants near (or below) a desired level. This problem can be conveniently accommodated in the optimal control framework for convection-diffusion equation. Some existing works ([27, 28, 29, 30]) focus on the stationary convection dominated optimal control problem. They used several stabilization methods to improve the approximation properties of the pure Galerkin discretization and to reduce the oscillatory behavior, e.g SUPG method in [27], stabilization on the Lagrangian functional method in [28], reduced basis (RB) technique in [29]. However to our best knowledge, there has been a lack of proper study for general time-dependent convection-diffusion optimal control problem.

The purpose of this paper is to extend the discontinuous Galerkin method with interior penalties in [17, 22] to time-dependent convection-diffusion optimal control problem. A semi-discrete time DG scheme for this problem is presented. The first difficulty for our problem is to derive the discretization of the co-state equation and the optimality conditions. We first establish the semi-discrete time DG scheme for the state equation, prove the stability and the existence of this scheme, then apply the theory of optimal control problem (see, [31]) to this scheme for deriving the discretization of the co-state equation and the optimality conditions. The DG scheme of state equation is complicated so that it is much more difficult to derive the discretized co-state equation, which is quite complicated. The complexity of the DG schemes of the state and the co-state equation also leads to the difficulties in deriving a priori error estimates and a posteriori error estimates later. To our knowledge, this paper appears to be the first trial to approximate convection-diffusion optimal control problem by using the Discontinuous Galerkin method with interior penalties.

The outline of the paper is as follows. In Section 2, we first briefly introduce convection-diffusion optimal control problem and optimality conditions. In Section 3, we give some definitions, then use discontinuous Galerkin method with interior penalties to construct a semi-discrete approximate scheme for convection-diffusion optimal control problem.

For this scheme, we prove the stability and the existence of the approximate solution. Then by the theory of optimal control problem, we present the semi-discrete optimality conditions. A priori error estimates are derived for both the state and the control approximation in Section 4. In Section 5, a posteriori error estimates are discussed for the case of an obstacle constraint under an assumption the velocity vector is incompressible. This assumption is needed to give stability bounds for the corresponding dual problem (see, Lemma 5.4) and not be satisfied by general convection dominated problems. Hence, we point out that the theoretical analysis of a posteriori error estimates here is not valid for general convection dominated problems. We will research on a posteriori error estimates for general convection dominated problems later.

2. Convection-diffusion Optimal Control Problem

Let Ω and Ω_U be bounded convex polygon domains in $R^n\,(n \leq 3)$ with Lipchitz boundary $\Gamma = \partial\Omega$ and $\partial\Omega_U$. In this paper, we adopt the standard notations for Sobolev spaces on Ω and its norms. In addition, c or C denotes a general positive constant independent of the mesh size h.

We shall take the space $W = L^2(0,T;V)$ with $V = H^1(\Omega)$, the control space $X = L^2(0,T;U)$ with $U = L^2(\Omega_U)$. The state space will be specified later. Let B be a bounded linear continuous operator from $L^2(0,T;U)$ to $L^2(0,T;L^2(\Omega))$. Let K be a closed convex set in $U = L^2(\Omega_U)$. Let $g(\cdot)$ be a convex functional which is continuously differential on $L^2(\Omega)$, and $h(\cdot)$ be a strictly convex continuously differential functional on U. We further assume that $h(u) \to +\infty$ as $\|u\|_U \to \infty$ and that $g(\cdot)$ is bounded below.

We are interested in the following convection-diffusion optimal control problem:

$$J(u) = \min_{u(t)\in K} \{\int_0^T (g(y) + h(u))dt\}, \tag{2.1}$$

subject to

$$\begin{cases} \frac{\partial y}{\partial t} - \nabla\cdot(a\nabla y) + \beta\cdot\nabla y + \alpha y = f + Bu, \qquad x\in\Omega,\ t\in(0,T], \\ y(x,0) = y_0(x), \qquad x\in\Omega, \\ y|_{\Gamma_-} = 0, \qquad t\in(0,T], \\ \big(a(x)\nabla y\big)\cdot\nu = 0, \qquad x\in\Gamma_+,\ t\in(0,T], \end{cases} \tag{2.2}$$

where $f(x,t) \in L^2(0,T;L^2(\Omega))$, $y_0(x) \in H^1(\Omega)$, and $a(x) = (a_{ij}(x))_{n\times n} \in (C^\infty(\bar\Omega))^{\,n\times n}$ such that there is a constant $a_0 > 0$ satisfying

$$\xi^T a(x)\xi \geq a_0|\xi|^2, \qquad \forall\, \xi \in R^n, \tag{2.3}$$

and $\beta(x) = (\beta_1(x), \beta_2(x), \cdots, \beta_n(x))^T$, ν is the outer normal vector to $\Gamma = \Gamma_- \cup \Gamma_+$,

$$\Gamma_- = \{x\in\Gamma:\ \beta(x)\cdot\nu(x) < 0\}, \qquad \Gamma_+ = \{x\in\Gamma:\ \beta(x)\cdot\nu(x) > 0\}.$$

Assuming that $\beta_i(x) \in C^1(\bar\Omega)$, $i = 1,2,\cdots,n$, $\alpha(x) \in C(\bar\Omega)$ and a constant $c_0 > 0$ satisfying

$$\alpha(x) - \frac{1}{2}div\beta(x) = c_0(x) > c_0. \tag{2.4}$$

Let $a(v,w)=\int_\Omega (a(x)\nabla v)\cdot\nabla w,\ \forall\ v,\ w\ \in H^1(\Omega);\ \ (v,w)_U=\int_{\Omega_U} vw,\ \forall\, v,\ w\ \in L^2(\Omega_U);$ $(f_1,f_2)=\int_\Omega f_1f_2,\ \forall\ f_1,\ f_2\ \in L^2(\Omega).$

It follows from the assumption on $a(x)$ that there are constants c and $C>0$ such that

$$a(v,\,v)\geq c\|v\|^2_{1,\Omega},\qquad |a(v,\,w)|\leq C|v|_{1,\Omega}\,|w|_{1,\Omega},\qquad \forall\, v,\ w\ \in\ H^1(\Omega).$$

The weak form of the convex optimal control problem reads:

$$J(u)=\min_{u(t)\in K}\{\int_0^T (g(y)+h(u))dt\},$$

where $y\in H^1(0,T;L^2(\Omega))\cap W,\ \ u\in L^2(0,T;L^2(\Omega_U)),\ u(t)\in K$ subject to

$$\begin{cases} (\dfrac{\partial y}{\partial t},w)+a(y,w)+(\beta\cdot\nabla y+\alpha y,w)=(f+Bu,w),\ \ t\in(0,T],\\ y(x,0)=y_0(x),\quad x\in\Omega,\\ y|_{\Gamma_-}=0,\quad t\in(0,T],\\ \big(a(x)\nabla y\big)\cdot\nu=0,\ \ x\ \in\Gamma_+,\, t\in(0,T], \end{cases}\tag{2.5}$$

for $w\in H^1(\Omega)$. It is well known (see e.g., [32]) that the above problem admits a unique solution y.

By the theory of optimal control problem (see, [31]), we can deduce that: the control problem (2.5) has a unique solution $(y,\ u)$, and that a pair $(y,\ u)$ is the solution iff there is a co-state $p\in H^1(0,T;L^2(\Omega))\cap W$ such that the triplet $(y,\ p,\ u)$ satisfies **the following optimality conditions:**

(2.6)

$$\begin{cases} (a)\quad (\frac{\partial y}{\partial t},w)+a(y,w)+(\beta\cdot\nabla y+\alpha y,w)=(f+Bu,w), & \forall\ w\in H^1(\Omega),\\ (b)\quad y(x,0)=y_0(x),\qquad y|_{\Gamma_-}=0,\qquad \big(a(x)\nabla y\big)\cdot\nu|_{\Gamma_+}=0, & \\ (c)\quad -(\frac{\partial p}{\partial t},q)+a(q,p)+(-\nabla\cdot(\beta p)+\alpha p,q)=(g'(y),q), & \forall\ q\in H^1(\Omega),\\ (d)\quad p(x,T)=0,\qquad p|_{\Gamma_+}=0,\qquad \big(a(x)\nabla p\big)\cdot\nu|_{\Gamma_-}=0, & \\ (e)\quad \int_0^T(h'(u)+B^*p,v-u)_U\,dt\geq 0,\qquad u(t)\in\ K,\forall\ v\in K, & \end{cases}$$

where B^* is the adjoint operator of B, g' and h' are the derivatives of g and h, which have been viewed as functions in $L^2(\Omega)$ and $L^2(0,T;L^2(\Omega_U))$, respectively.

3. Semi-discrete DG approximation

3.1. Preliminaries. Let us consider the discontinuous Galerkin finite element approximation of the control problem (2.5). Define $T^h=\{\tau_1,\,\tau_2,\,\cdots,\tau_{N_h}\}$ be a non-degenerate quasi-uniform subdivision of Ω. Each element has at most one face on Γ, and two neighboring elements have either only one common vertex or a whole edge ($n=2$) or face ($n=3$). Let $h_j=diam(\tau_j)$ and $h_\tau=max\{h_j\}$, $j=1,\cdots,N_h$. Here, the non-degeneracy requirement is that there exists a constant $\rho>0$ such that each τ_j contains a ball of radius ρh_j. And the quasi-uniformity requirement is that there is a constant $\gamma>0$ such that $h/h_j \leq \gamma$ for all $j \in 1, \cdots, N_h$. For an element $\tau \in T^h$, we denote $\partial\tau$ is the union of open faces of τ. Let $x \in \partial\tau$ and suppose that $n_\tau(x)$ denote the unit outward normal vector to $\partial\tau$ at x. With these conventions, we define the inflow and outflow parts of $\partial\tau$, respectively, by

$$\partial\tau_- = \{x \in \partial\tau : \ \beta(x) \cdot n_\tau(x) < 0\}, \qquad \partial\tau_+ = \{x \in \partial\tau : \ \beta(x) \cdot n_\tau(x) > 0\}.$$

For an element $\tau \in T^h$ and $v \in H^1(\tau)$, we denote by v^+ the interior trace of v on $\partial\tau$, i.e. the trace taken from within τ. Now considering an element τ such that the set of $\partial\tau \setminus \Gamma_-$ is nonempty; then for each $x \in \partial\tau \setminus \Gamma_-$ there exists a unique element τ', depending on the choice of x, such that $x \in \partial\tau'_+$.

Now suppose that $v \in H^1(\tau)$ for each $\tau \in T^h$. If $\partial\tau \setminus \Gamma_-$ is nonempty for some $\tau \in T^h$, then we can also define the outer trace v^- of v on $\partial\tau \setminus \Gamma_-$ relative to τ as the inner trace v^+ relative to those elements τ' for which $\partial\tau'_+$ has intersection with $\partial\tau \setminus \Gamma_-$ of positive $(d-1)$-dimensional measure. We also introduce the average and jump of such v across $\partial\tau \setminus \Gamma_-$

$$\{v\} = \frac{1}{2}(v^+ + v^-), \qquad [v] = v^+ - v^-.$$

Let $r \geq 1$ be a positive integer. The finite element space associated with T^h is taken to be

$$V^h = \{v \in L^2(\Omega) : v|_\tau \in \ P_r(\tau), \ \forall \ \tau \in T^h\}, \tag{3.1}$$

where $P_r(\tau)$ denotes the set of polynomials of degree less than or equal to r on τ. With each edge (or face) e_k, we associate a unit normal vector ν_k. For $k > P_h, \nu_k$ is taken to be the unit outward vector normal to $\partial\Omega$. The norms associated with this space are the following "broken" norms for positive integer m ([33]):

$$|||\phi|||^2 = \sum_{j=1}^{N_h} \|\phi\|^2_{0,\,\tau_j}, \qquad |||\phi|||^2_{L^2((\alpha,\,\beta);\,L^2(\Omega))} = \int_\alpha^\beta |||\phi(\cdot,t)|||^2 dt,$$
$$|||\phi|||^2_m = \sum_{j=1}^{N_h} \|\phi\|^2_{m,\,\tau_j}, \qquad |||\phi|||^2_{L^2((\alpha,\,\beta);\,H^m(\Omega))} = \int_\alpha^\beta |||\phi(\cdot,t)|||^2_m dt.$$

It is easy to see that $V^h \not\subset V$. For later use, we define the space $Y^h = H^1(0,T,V^h)$.

We denote the edges (or faces) of the elements by $\{e_1, e_2, \cdots, e_{P_h}, e_{P_{h+1}}, \cdots, e_{M_h}\}$, where $e_k \subset \Omega, \ 1 \leq k \leq P_h$, and $e_k \subset \partial\Omega, \ P_{h+1} \leq k \leq M_h$. The interior penalty term is defined as

$$J_0^\sigma(\phi,\psi) = \sum_{k=1}^{P_h} \frac{\sigma_k}{|h_{e_k}|} \int_{e_k} [\phi][\psi], \tag{3.2}$$

where $|h_{e_k}|$ denotes the measure of e_k and σ_k is a real nonnegative constant associated to the interior edge e_k, which is bounded below by $\sigma_0 > 0$, above by σ^*.

For each $\tau_j \in T^h$, we denote $\partial\tau_j$ is an edge (or a face) of τ_j. By regular subdivision T^h of Ω, we hold the following approximation properties([22]). There exists a constant C depending on r, ρ, γ such that the local inverse inequalities

$$\|\phi\|^2_{0,\partial\tau_j} \leq Ch_j^{-1}\|\phi\|^2_{0,\tau_j}, \qquad \|\frac{\partial\phi}{\partial\nu_k}\|^2_{0,\partial\tau_j} \leq Ch_j^{-1}\|\nabla\phi\|^2_{0,\tau_j}, \tag{3.3}$$

are valid for $\phi \in P_r(\tau)$. Directly from the above inverse inequalities, there exists a constant C depending only on ρ, γ such that

$$\sum_{k=1}^{P_h} |h_{e_k}| \|\{\frac{\partial \phi}{\partial \nu_k}\}\|_{0,e_k}^2 \leq C_1 |||\nabla \phi|||_0^2, \quad \forall \phi \in V^h. \tag{3.4}$$

This inequality will be used often later.

Let T_U^h be a partition of Ω into disjoint regular n-simplices τ_U. Each element has at most one face on Γ, and two neighbor elements have either only one common vertex or a whole edge or face. Let h_{τ_U} denote the maximum diameter of the element τ_U in T_U^h. Associated with T_U^h is another finite element space

$$U^h = \{v \in L^2(\Omega_U) : v|_{\tau_U} \in P_m(\tau_U), \ \forall \ \tau_U \in T_U^h\}, \tag{3.5}$$

where $P_m(\tau_U)$ denotes the set of polynomials of degree less than or equal to $m \geq 0$ on τ_U. The definitions of "broken" norms for U^h are similar to that of V^h. Let $X^h = L^2(0, T; U^h)$. It is easy to see that $U^h \subset U, \ X^h \subset X$. Let K^h be an approximation of K. For ease of exposition, we assume that K^h is a closed convex set in U^h and $K^h \subset U^h \cap K$. More complicated cases can be considered following the approach in [10].

Note that in general the sizes of the elements in T_U^h are smaller than those in T^h in computations. Therefore, we assume that $h_{\tau_U}/h_\tau \leq C$ in this paper.

3.2. DG Scheme for the state equation. Using the above notations and definitions, we present the following: semi-discrete DG method of convection-diffusion optimal control problem is $(QCP)^h$

$$\min_{u_h(t) \in K^h} \{ \int_0^T (g(y_h) + h(u_h))\, dt\}, \tag{3.6}$$

with $y_h \in Y^h = H^1(0, T; V^h)$ subject to

$$\begin{cases} (a) \ (\frac{\partial y_h}{\partial t}, w_h) + \sum\limits_{j=1}^{N_h} \int_{\tau_j} a(x) \nabla y_h \cdot \nabla w_h - \sum\limits_{k=1}^{P_h} \int_{e_k} \{a(x) \nabla y_h \cdot \nu_k\}[w_h] \\ \quad + \sum\limits_{k=1}^{P_h} \int_{e_k} \{a(x) \nabla w_h \cdot \nu_k\}[y_h] + J_0^\sigma(y_h, w_h) + (\beta \cdot \nabla y_h + \alpha y_h, w_h) \\ \quad - \sum\limits_{\tau} \int_{\partial\tau_- \backslash \Gamma_-} \beta \cdot n_\tau [y_h] w_h^+ - \sum\limits_{\tau} \int_{\partial\tau_- \cap \Gamma_-} \beta \cdot n_\tau y_h^+ w_h^+ \\ \quad = (f + Bu_h, w_h), \qquad \forall \ w_h \in V^h, \\ (b) \ y_h(x, 0) = y_0^h(x), \qquad y_h|_{\Gamma_-} = 0, \qquad \big(a(x) \nabla y_h\big) \cdot \nu|_{\Gamma_+} = 0, \end{cases} \tag{3.7}$$

where $y_0^h \in V^h$ is an approximation of $y_0(x)$.

Then, we introduce a nonsymmetric bilinear form: $\quad \forall \ \phi, \ \psi \in V^h$,

$$\begin{aligned} A(a(x); \phi, \psi) = & \sum_{j=1}^{N_h} \int_{\tau_j} a(x) \nabla \phi \cdot \nabla \psi - \sum_{k=1}^{P_h} \int_{e_k} \{a(x) \nabla \phi \cdot \nu_k\}[\psi] \\ & + \sum_{k=1}^{P_h} \int_{e_k} \{a(x) \nabla \psi \cdot \nu_k\}[\phi]. \end{aligned} \tag{3.8}$$

Suppose that $y_h,\ v_h \in V^h$ and $\forall\ \tau \in\ T^h$, we define two bilinear forms as follow

$$l(y_h,\, v_h) = (\beta\cdot\nabla y_h + \alpha y_h, v_h)_\tau - \int_{\partial\tau_-\backslash\Gamma_-} \beta\cdot n_\tau [y_h] v_h^+ - \int_{\partial\tau_-\cap\Gamma_-} \beta\cdot n_\tau y_h^+ v_h^+, \tag{3.9a}$$

$$L(y_h,\, v_h) = \sum_\tau l(y_h,\, v_h). \tag{3.9b}$$

For ease of exposition, in this paper we introduce an inner product and a corresponding norm on edge (or face) e of an element τ as follow

$$(w,\, v)_e = \int_e |\beta\cdot\nu| w\, v\, ds, \qquad \|v\|_e^2 = \int_e |\beta\cdot\nu| v^2\, ds.$$

Now we turn to prove a stability lemma for the state equation, which is useful in the rest of the paper.

Lemma 3.1. *Suppose that there exits a positive constant c_0 such that (2.4) holds. Then y_h of (3.7) obeys the following bound* $\forall\ t \in\ (0, T]$

$$\begin{aligned} &|||y_h(t)|||^2 + 2a_0\int_0^t |||\nabla y_h|||^2\, dt + 2\int_0^t J_0^\sigma(y_h,\, y_h) dt \\ &+ \int_0^t \sum_\tau \big\{c_0\|y_h\|_\tau^2 + \|y_h^+\|^2_{\partial\tau_-\cap\Gamma_-} + \|[y_h]\|^2_{\partial\tau_-\backslash\Gamma_-} + \|y_h^+\|^2_{\partial\tau_+\cap\Gamma_+}\big\}\, dt \\ &\le |||y_0^h(x)|||^2 + \frac{1}{c_0}\int_0^t \big\{|||f|||^2 + |||Bu_h|||^2\big\}\, dt. \end{aligned} \tag{3.10}$$

Proof. The proof is similar to that of Lemma 2.4 in [34]. Taking $w_h = y_h$ in (3.7a), this gives

$$(\frac{\partial y_h}{\partial t}, y_h) + A(a(x); y_h, y_h) + J_0^\sigma(y_h, y_h) + L(y_h,\, y_h) = (f + Bu_h, y_h). \tag{3.11}$$

Upon partial integration, we have

(3.12)

$$\begin{aligned} &\text{the left-hand side of (3.11)} \\ &= \frac{d}{2dt}|||y_h|||^2 + \sum_{j=1}^{N_h}\int_{\tau_j} (a(x)\nabla y_h,\, \nabla y_h)_{\tau_j} + J_0^\sigma(y_h, y_h) + \sum_\tau\{\int_\tau (\alpha - \frac{1}{2}div\beta)|y_h|^2 \\ &+ \frac{1}{2}\int_{\partial\tau}(\beta\cdot n_\tau)|y_h^+|^2\, ds - \int_{\partial\tau_-\backslash\Gamma_-}(\beta\cdot n_\tau)[y_h]y_h^+\, ds - \int_{\partial\tau_-\cap\Gamma_-}(\beta\cdot n_\tau)|y_h^+|^2\, ds\}. \end{aligned}$$

The last three terms in (3.12) can be rewritten as

$$\begin{aligned} &\frac{1}{2}\sum_\tau \int_{\partial\tau_-\cap\Gamma_-} -(\beta\cdot n_\tau)|y_h^+|^2\, ds + \frac{1}{2}\sum_\tau\int_{\partial\tau_-\backslash\Gamma_-} -(\beta\cdot n_\tau)|y_h^+ - y_h^-|^2\, ds \\ &+\frac{1}{2}\sum_\tau\int_{\partial\tau_+\cap\Gamma_+}(\beta\cdot n_\tau)|y_h^+|^2\, ds. \end{aligned} \tag{3.13}$$

Using (3.13) in (3.12) yields

$$\begin{aligned} &\text{the left-hand side of (3.11)} \\ &\geq \frac{d}{2dt}|||y_h|||^2 + a_0|||\nabla y_h|||^2 + J_0^\sigma(y_h, y_h) + c_0 \sum_\tau \|y_h\|_\tau^2 \\ &+\frac{1}{2}\sum_\tau \|y_h^+\|^2_{\partial\tau_-\cap\Gamma_-} + \frac{1}{2}\sum_\tau \|y_h^+ - y_h^-\|^2_{\partial\tau_-\setminus\Gamma_-} + \frac{1}{2}\sum_\tau \|y_h^+\|^2_{\partial\tau_+\cap\Gamma_+}. \end{aligned} \tag{3.14}$$

Now we bound the right-hand side in (3.11):

$$|(f + Bu_h, y_h)| \leq \frac{c_0}{2}\sum_\tau \|y_h\|_\tau^2 + \frac{1}{c_0}\{|||f|||^2 + |||Bu_h|||^2\}. \tag{3.15}$$

Inserting (3.14),(3.15) into (3.11) and integrating time from 0 to t lead to (3.10). □

By Lemma 3.1, we can prove the following existence theorem.

Theorem 3.1. (Existence Theorem) *Let $J(\cdot)$ be a continuous functional in U. Suppose that $h(u) \to +\infty$ as $\|u\|_U \to \infty$. Then there exists at least one solution for the minimization problem (3.6).*

Proof. Let $u_h^n \in K^h$ be a minimization sequence. Then it is clear that u_h^n are bounded in $L^2(0,T;L^2(\Omega_U))$. Thus there is a subsequence u_h^n such that u_h^n converge to u_h^* weakly in $L^2(0,T;L^2(\Omega_U))$.

For the subsequence u_h^n, we have

$$\begin{aligned} &(\tfrac{\partial}{\partial t}y(u_h^n), w_h) + A(a(x); y(u_h^n), w_h) + J_0^\sigma(y(u_h^n), w_h) + L(y(u_h^n),\ w_h) \\ &= (f + Bu_h^n, w_h), \qquad \forall\ w_h \in V^h. \end{aligned} \tag{3.16}$$

By Lemma 3.1, we know that $|||y(u_h^n)|||_{L^2(0,T;H^1(\Omega))}$ is bounded. Thus

$$\begin{aligned} y(u_h^n) &\to y_h^* \qquad \text{weakly in} \qquad L^2(0,T;H^1(\Omega)), \\ y(u_h^n) &\to y_h^* \qquad \text{strongly in} \qquad L^2(0,T;L^2(\Omega)). \end{aligned}$$

By trace theorem: $H^1(\Omega) \hookrightarrow H^{1/2}(\Gamma)$, we have

$$\|u\|_{H^{1/2}(\Gamma)} \leq c\|u\|_{H^1(\Omega)} \text{ and } \|\frac{\partial u}{\partial \nu}\|_{H^{-1/2}(\Gamma)} \leq c\|u\|_{H^1(\Omega)}, \ \forall\ u \in\ H^1(\Omega), \tag{3.17}$$

Hence, $|||y(u_h^n)|||_{L^2(0,T;H^{1/2}(\Gamma))}$ and $|||\frac{\partial y(u_h^n)}{\partial \nu}|||_{H^{-1/2}(\Gamma)}$ are bounded. Thus

$$\begin{aligned} y(u_h^n) &\to y_h^* \qquad \text{weakly in} \qquad L^2(0,T;H^{1/2}(\Gamma)), \\ \frac{\partial y(u_h^n)}{\partial \nu} &\to \frac{\partial y_h^*}{\partial \nu} \qquad \text{weakly in} \qquad L^2(0,T;H^{-1/2}(\Gamma)). \end{aligned}$$

So, we have

$$(\frac{\partial}{\partial t}y_h^*, w_h) + A(a(x); y_h^*, w_h) + J_0^\sigma(y_h^*, w_h) + L(y_h^*),\ w_h) = (f + Bu_h^*, w_h). \tag{3.18a}$$

Since $g(\cdot)$ be a convex functional on space $L^2(\Omega)$ and $h(\cdot)$ be a strictly convex functional on U, we have

$$\int_0^T (g(y_h^*) + h(u_h^*))\,dt\ \leq \underline{\lim}\{\int_0^T (g(y_h^n) + h(u_h^n))\,dt\}. \tag{3.18b}$$

Thus $(y_h^*,\ u_h^*)$ is a solution of (3.7). □

3.3. Optimality conditions. Supposing that $p_h,\ q_h \in V^h$ and $\forall\ \tau \in\ T^h$, we define other two bilinear forms as follow

$$l^*(p_h,\, q_h) = (-\nabla\cdot(\beta p_h)+\alpha p_h, q_h)_\tau + \int_{\partial\tau_+\setminus\Gamma_+} \beta\cdot n_\tau [p_h] q_h^+ + \int_{\partial\tau_+\cap\Gamma_+} \beta\cdot n_\tau p_h^+ q_h^+, \tag{3.19a}$$

$$L^*(p_h,\, q_h) = \sum_\tau l^*(p_h,\, q_h). \tag{3.19b}$$

By the theory of optimal control problem (see, [31]), we can deduce that the control problem $(QCP)^h$ has a unique solution $(y_h,\, u_h)$ and that a pair $(y_h,\ u_h) \in Y^h \times X^h$ is the solution of $(QCP)^h$ iff there is a co-state $p_h \in Y^h$ such that the triplet $(y_h,\ p_h,\ u_h) \in Y^h \times Y^h \times X^h$ satisfies the following optimality conditions: $(QCP-OPT)^h$

$$\begin{cases} (a) & (\frac{\partial y_h}{\partial t}, w_h) + A(a(x); y_h, w_h) + J_0^\sigma(y_h, w_h) + L(y_h,\, w_h) \\ & \qquad = (f + Bu_h, w_h),\ \ \forall\ w_h \in V^h, \\ (b) & y_h(x,0) = y_0^h(x), \qquad y_h|_{\Gamma_-} = 0, \qquad \big(a(x)\nabla y_h\big)\cdot\nu|_{\Gamma_+} = 0, \\ (c) & -(\frac{\partial p_h}{\partial t}, q_h) + A(a(x); q_h, p_h) + J_0^\sigma(p_h, q_h) + L^*(p_h,\, q_h) \\ & \qquad = (g'(y_h), q_h),\ \ \forall\ q_h \in V^h, \\ (d) & p_h(x,T) = 0, \qquad p_h|_{\Gamma_+} = 0, \qquad \big(a(x)\nabla p_h\big)\cdot\nu|_{\Gamma_-} = 0, \\ (e) & \int_0^T (h'(u_h) + B^* p_h, v_h - u_h)_U\, dt \geq 0,\ \ \forall\ v_h \in K^h. \end{cases} \tag{3.20}$$

Similarly to Lemma 3.1, we can get the following stability lemma for the co-state equation.

Lemma 3.2. *Suppose that there exits a positive constant c_0 such that (2.4) holds. Then p_h of (3.20c) obeys the following bound*

$$\begin{aligned} &|||p_h(t)|||^2 + 2a_0 \int_t^T |||\nabla p_h|||^2\, dt + 2\int_t^T J_0^\sigma(p_h,\, p_h) dt \\ &+ \int_t^T \sum_\tau \big\{c_0\|p_h\|_\tau^2 + \|p_h^+\|_{\partial\tau_-\cap\Gamma_-}^2 + \|[p_h]\|_{\partial\tau_+\setminus\Gamma_+}^2 + \|p_h^+\|_{\partial\tau_+\cap\Gamma_+}^2\big\}\, dt \\ &\leq \frac{1}{c_0}\int_t^T |||g'(y_h)|||^2\, dt. \end{aligned} \tag{3.21}$$

The proof of (3.21) is analogous to that of (3.10) by using $p_h(x,\, T) = 0$.

4. A priori error estimates

In this section, we shall derive a priori error estimates for the semi-discrete DG schemes (3.20). For ease of exposition, we simply write $L^2(0,T;L^2(\Omega_U))$ as $L^2(L^2(\Omega_U))$, $L^2(0,T;L^2(\Omega))$ as $L^2(L^2)$, $L^2(0,T;H^1(\Omega_U))$ as $L^2(H^1(\Omega_U))$, and $L^2(0,T;H^1(\Omega))$ as $L^2(H^1)$, etc. in the following contents of the paper.

We shall assume that the following convexity conditions:

$$(h'(u) - h'(v), u - v) \geq c\|u - v\|_{0,\Omega_U}^2, \qquad \forall\, u,\ v \in L^2(\Omega_U), \tag{4.1a}$$

that is to say $h(\cdot)$ is uniformly convex.

Noting that g is convex, it is easy to see that

$$(g'(u)-g'(v),u-v)\ge 0,\quad \forall\, u,\ v\in H^1(\Omega). \tag{4.1b}$$

Also, we have that

$$|(Bv,w)|=|(v,B'w)|\le c\|v\|_{0,\Omega_U}\|w\|_{0,\Omega},\quad \forall\, v\in L^2(\Omega_U),\ w\in H^1(\Omega), \tag{4.2}$$

because that B is a bounded linear operator.

Let

$$J_h'(u)(v-u)=\int_0^T (h'(u)+B^*p_h(u),v-u)_U\,dt,\ \ \forall\ v\in K,$$

where $p_h(u)\in Y^h$ is the solution of the system:

$$\begin{cases} (a)\quad (\frac{\partial y_h(u)}{\partial t},w_h)+A(a(x);y_h(u),w_h)+J_0^\sigma(y_h(u),w_h)+L(y_h(u),\,w_h)\\ \qquad\qquad =(f+Bu,w_h),\quad \forall\ w_h\ \in V^h,\\ (b)\quad y_h(u)(x,0)=y_0^h(x),\qquad y_h(u)|_{\Gamma_-}=0,\quad \{a(x)\nabla y_h(u)\}\cdot\nu|_{\Gamma_+}=0,\\ (c)\quad -(\frac{\partial p_h(u)}{\partial t},q_h)+A(a(x);q_h,p_h(u))+J_0^\sigma(p_h(u),q_h)+L^*(p_h(u),\,q_h)\\ \qquad\qquad =(g'(y_h(u)),q_h),\quad \forall\ q_h\ \in V^h,\\ (d)\quad p_h(u)(x,T)=0,\qquad p_h(u)|_{\Gamma_+}=0,\quad \{a(x)\nabla p_h(u)\}\cdot\nu|_{\Gamma_-}=0. \end{cases} \tag{4.3}$$

To derive a priori estimates, we need prove the following three lemmas.

Lemma 4.1 *If $h(\cdot)$ is uniformly convex, and $g(\cdot)$ is convex, then*

$$J_h'(v)(v-u)-J_h'(u)(v-u)\ge c\|v-u\|^2_{L^2(L^2(\Omega_U))}. \tag{4.4}$$

Proof. Note that

$$\begin{aligned} &J_h'(v)(v-u)-J_h'(u)(v-u)\\ &=\int_0^T (h'(v)-h'(u),v-u)_U\,dt+\int_0^T (B^*p_h(v)-B^*p_h(u),v-u)_U\,dt. \end{aligned} \tag{4.5}$$

Moreover, it follows from (4.3) that

$$\int_0^T (B^*p_h(v)-B^*p_h(u),v-u)_U\,dt=\int_0^T (g'(y_h(v))-g'(y_h(u)),y_h(v)-y_h(u))\,dt. \tag{4.6}$$

Noting that $h(\cdot)$ is uniformly convex and $g(\cdot)$ is convex, (4.5) and (4.6) imply that

$$J_h'(v)(v-u)-J_h'(u)(v-u)\ge\int_0^T (h'(v)-h'(u),v-u)_U\,dt\ge c\|v-u\|^2_{L^2(L^2(\Omega_U))}.$$

This proves (4.4). □

Lemma 4.2 *Let (y,p,u) and (y_h,p_h,u_h) be the solutions of optimality conditions (2.6) and semi-discrete DG optimality conditions (3.20), respectively. Assume that $u\in L^2(0,T;H^1(\Omega_U))$, $p\in L^2(0,T;H^1(\Omega))$, $K^h\subset U^h\cap K$, $h'(\cdot)$ is Lipschitz continuous,*

$u_I \in K^h$, where u_I is the standard Lagrange interpolation of u. Moreover, assume that all conditions of Lemma 4.1 exist. Then

$$|||u-u_h|||^2_{L^2(L^2(\Omega_U))} \leq Ch_U^2\{|||u|||^2_{L^2(H^1(\Omega_U))}+|||p|||^2_{L^2(H^1)}\}+C|||p_h(u)-p|||^2_{L^2(L^2)}, \tag{4.7}$$

where $p_h(u) \in Y^h$ is the solution of the system (4.3).

Proof. It follows from (2.6), (3.20) and Lemma 4.1 that

$$\begin{aligned} c|||u-u_h|||^2_{L^2(L^2(\Omega_U))} &\leq J_h'(u)(u-u_h) - J_h'(u_h)(u-u_h) \\ &\leq \int_0^T (B^*p - B^*p_h(u), u_h - u)_U\, dt + \int_0^T (h'(u_h) + B^*p_h, u_I - u)_U\, dt. \end{aligned} \tag{4.8}$$

Note that

$$\begin{aligned} &\int_0^T (h'(u_h) + B^*p_h, u_I - u)_U\, dt \\ &\leq \int_0^T (h'(u) + B^*p, u_I - u)_U\, dt + C(\delta)|||u_I - u|||^2_{L^2(L^2(\Omega_U))} \\ &\quad + C\delta|||h'(u_h) - h'(u)|||^2_{L^2(L^2(\Omega_U))} + C\delta|||B^*p_h - B^*p_h(u)|||^2_{L^2(L^2(\Omega_U))} \\ &\quad + C\delta|||B^*p_h(u) - B^*p|||^2_{L^2(L^2(\Omega_U))}, \end{aligned} \tag{4.9}$$

where δ is an arbitrary small positive constant. Moreover, we have

$$\int_0^T (h'(u) + B^*p, u_I - u)_U\, dt \leq Ch_U^2\{|||u|||^2_{L^2(H^1(\Omega_U))} + |||p|||^2_{L^2(H^1)}\}, \tag{4.10}$$

and

$$|||u_I - u|||_{L^2(L^2(\Omega_U))} \leq Ch_U|||u|||_{L^2(H^1(\Omega_U))}. \tag{4.11}$$

Then it follows from (4.8)-(4.11) that

$$\begin{aligned} |||u-u_h|||^2_{L^2(L^2(\Omega_U))} &\leq Ch_U^2(|||u|||^2_{L^2(H^1(\Omega_U))} + |||p|||^2_{L^2(H^1)}) + C\delta|||p_h(u) - p|||^2_{L^2(L^2)} \\ &\quad + C\delta|||u-u_h|||^2_{L^2(L^2(\Omega_U))} + C\delta|||p_h - p_h(u)|||^2_{L^2(L^2)}. \end{aligned} \tag{4.12}$$

Furthermore, from (3.20), (4.3), Lemma 3.1 and Lemma 3.2, we can deduce that

$$|||p_h - p_h(u)|||_{L^2(L^2)} \leq C|||y_h - y_h(u)|||_{L^2(L^2)} \leq C|||u-u_h|||_{L^2(L^2(\Omega_U))}. \tag{4.13}$$

Then (4.12) and (4.13) prove (4.7). □

Lemma 4.3. *Let (y, p, u) and $(y_h(u), p_h(u))$ be the solutions of (2.6) and (4.3), respectively. Assume that $g'(\cdot)$ is Lipschitz continuous, $y,\ p \in L^2(0,T;H^2(\Omega))\cap H^1(0,T;H^1(\Omega))$. Then*

$$\begin{aligned} &|||y_h(u) - y|||_{L^2(H^1)} + |||p_h(u) - p|||_{L^2(H^1)} \\ &\leq Ch\{|||y|||_{L^2(H^2)} + |||y_t|||_{L^2(H^1)} + |||p|||_{L^2(H^2)} + |||p_t|||_{L^2(H^1)}\}. \end{aligned} \tag{4.14}$$

Proof. Note that $y_h(u)$ is the semi-DG finite element solution of y. Then, from (4.3a) and (2.6a) we have

$$\begin{aligned}&(\frac{\partial}{\partial t}(y_h(u)-y),w_h)+A(a(x);y_h(u)-y,w_h)+J_0^\sigma(y_h(u)-y,w_h)\\&+L(y_h(u)-y,\,w_h)=0,\qquad \forall\ w_h\in\ V^h.\end{aligned}\tag{4.15}$$

Let $y_h(u)-y=\theta-\xi$, where $\theta=y_h(u)-y_I$, $\xi=y-y_I$ and y_I is the standard Lagrange interpolation of y. Taking $w_h=\theta$ in (4.15) and noting that $[\xi]\equiv 0$ on the interior edges ${e_k}_{k=1}^{P_h}$, we can obtain

$$\begin{aligned}&(\frac{\partial\theta}{\partial t},\theta)+A(a(x);\theta,\theta)+J_0^\sigma(\theta,\theta)+L(\theta,\,\theta)\\&=(\frac{\partial\xi}{\partial t},\theta)+(a(x)\nabla\xi,\nabla\theta)-\sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla\xi\cdot\nu_k\}[\theta]\\&+\sum_\tau\big\{(\beta\cdot\nabla\xi+\alpha\xi,\theta)_\tau-\int_{\partial\tau_-\cap\Gamma_-}\beta\cdot n_\tau\xi^+\theta^+\big\}.\end{aligned}\tag{4.16}$$

Similar to the inequality (3.14) in the proof of Lemma 3.1, it is easy to see that

$$\begin{aligned}&\text{the left-hand side of (4.16)}\\&\geq\frac{d}{2dt}|||\theta|||^2+a_0|||\nabla\theta|||^2+J_0^\sigma(\theta,\theta)+c_0|||\theta|||^2\\&+\frac12\sum_\tau\|\theta^+\|^2_{\partial\tau_-\cap\Gamma_-}+\frac12\sum_\tau\|[\theta]\|^2_{\partial\tau_-\backslash\Gamma_-}+\frac12\sum_\tau\|\theta^+\|^2_{\partial\tau_+\cap\Gamma_+}.\end{aligned}\tag{4.17}$$

Now we bound the terms on the right-hand side of (4.16).

(I)

$$|(\frac{\partial\xi}{\partial t},\theta)|\leq C|||\frac{\partial\xi}{\partial t}|||^2+\frac{c_0}{4}|||\theta|||^2.\tag{4.18}$$

(II)

$$\|(a(x)\nabla\xi,\nabla\theta)-\sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla\xi\cdot\nu_k\}[\theta]|\leq C|||\nabla\xi|||^2+\frac{a_0}{2}|||\nabla\theta|||^2+\frac12J_0^\sigma(\theta,\theta),\tag{4.19}$$

where we used the inequality (3.4).

(III)

$$\begin{aligned}&|\sum_\tau\big\{(\beta\cdot\nabla\xi+\alpha\xi,\theta)_\tau-\int_{\partial\tau_-\cap\Gamma_-}\beta\cdot n_\tau\xi^+\theta^+\big\}|\\&\leq C\{|||\nabla\xi|||^2+|||\xi|||^2\}+\frac{c_0}{4}|||\theta|||^2+C\sum_\tau\|\xi^+\|^2_{\partial\tau_-\cap\Gamma_-}+\frac14\sum_\tau\|\theta^+\|^2_{\partial\tau_-\cap\Gamma_-}.\end{aligned}\tag{4.20}$$

Combining (4.17)-(4.20) together, integrating time from 0 to T, taking $\theta(0)=0$ and by the trace theorem and Lagrange interpolant approximation property, we can derive

$$|||\theta|||_{L^2(H^1)}\leq Ch\{|||y|||_{L^2(H^2)}+|||y_t|||_{L^2(H^1)}\}.\tag{4.21}$$

Hence, we have

$$|||y_h(u)-y|||_{L^2(H^1)} \leq |||\theta|||_{L^2(H^1)}+|||\xi|||_{L^2(H^1)} \leq Ch\{|||y|||_{L^2(H^2)}+|||y_t|||_{L^2(H^1)}\}. \tag{4.22}$$

By (4.3c) and (2.6c) and Lemma 3.2, we can deduce similarly

$$|||p_h(u) - p|||_{L^2(H^1)} \leq Ch\{|||p|||_{L^2(H^2)} + |||p_t|||_{L^2(H^1)}\}. \tag{4.23}$$

Then (4.14) follows from (4.21) and (4.22). □

By Lemma 4.1 and 4.3, we can derive the following theorem for a priori estimates.

Theorem 4.1 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of optimality conditions (2.6) and semi-discrete DG optimality conditions (3.19), respectively. Assume that all conditions of Lemmas 4.1-4.3 are valid. Then*

$$\begin{aligned}
&|||u - u_h|||_{L^2(L^2(\Omega_U))} + |||y - y_h|||_{L^2(H^1)} + |||p - p_h|||_{L^2(H^1)} \\
&\leq Ch_U\{|||u|||_{L^2(H^1(\Omega_U))} + |||p|||_{L^2(H^1)}\} \\
&\quad +Ch\{|||y|||_{L^2(H^2)} + |||y_t|||_{L^2(H^1)} + |||p|||_{L^2(H^2)} + |||p_t|||_{L^2(H^1)}\}.
\end{aligned} \tag{4.24}$$

Proof. It follows from (4.7) and (4.23) that

$$\begin{aligned}
|||u - u_h|||_{L^2(L^2(\Omega_U))} &\leq Ch_U\{|||u|||_{L^2(H^1(\Omega_U))} + |||p|||_{L^2(H^1)}\} \\
&\quad +Ch\{|||p|||_{L^2(H^2)} + |||p_t|||_{L^2(H^1)}\}.
\end{aligned} \tag{4.25}$$

Moreover, it follows from (4.13), (4.14) and (4.25) that

$$\begin{aligned}
&|||y - y_h|||_{L^2(H^1)} + |||p - p_h|||_{L^2(H^1)} \\
&\leq Ch_U\{|||u|||_{L^2(H^1(\Omega_U))} + |||p|||_{L^2(H^1)}\} \\
&\quad +Ch\{|||y|||_{L^2(H^2)} + |||y_t|||_{L^2(H^1)} + |||p|||_{L^2(H^2)} + |||p_t|||_{L^2(H^1)}\}.
\end{aligned} \tag{4.26}$$

Then (4.24) follows from (4.25) and (4.26). □

5. A Posteriori Error Estimates

In this paper, we consider a posteriori error estimates only for the case of an obstacle constraint and assumption that the velocity vector $\beta(x)$ is incompressible i.e. $div\beta(x) = 0,\ \forall\, x \in \Omega$.

Remark 5.1. The assumption of the velocity vector $\beta(x)$ incompressible is needed to give stability bounds for the corresponding dual problem (cf. Lemma 5.4). For the case of compressible $\beta(x)$, similar stability estimates have been derived for a system of convection-diffusion problems in [39].

We assume that the constraint on the control is an obstacle such that

$$K = \{v \in X = L^2(0, T; L^2(\Omega_U)) : \ v \geq d,\ \ a.e. \text{ in } \Omega_U \times (0, T]\}, \tag{5.1}$$

where d is a constant. This obstacle constraint is met most frequently in practical application. We define the coincidence set (contact set) $\Omega_U^-(t)$ and the non-coincidence set (non-contact set) $\Omega_U^+(t)$ as follows:

$$\Omega_U^-(t) := \{x \in \Omega_U : u(x, t) = d\}, \qquad \Omega_U^+(t) := \{x \in \Omega_U : u(x, t) > d\}. \tag{5.2}$$

The convexity assumptions (4.1) for $h(\cdot)$ and $g(\cdot)$ will still be used.

Let

$$K^h = \{v \in U^h : \ v \geq d \text{ in } \Omega_U \times (0, T]\}. \tag{5.3}$$

Hence, we have that $K^h \subset K$.

It can be seen that the inequality in (2.6) is now equivalent to the followings:

$$h'(u) + B^*p \geq 0, \quad u \geq d, \qquad (h'(u) + B^*p)(u - d) = 0, \quad a.e. \text{ in } \Omega_U \times (0, T]. \tag{5.4}$$

In order to derive sharper a posteriori error estimate, we divide Ω_U into the following three subsets:

$$\Omega_d^- = \{x \in \Omega_U : B^*p_h(x,t) \leq -h'(d)\},$$

$$\Omega_d = \{x \in \Omega_U : B^*p_h(x,t) > -h'(d),\ u_h = d\},$$

$$\Omega_d^+ = \{x \in \Omega_U : B^*p_h(x,t) > -h'(d),\ u_h > d\}.$$

Then, it is easy to see that above three subsets are not overlapped each other, and

$$\bar{\Omega}_U = \bar{\Omega}_d^- \cup \bar{\Omega}_d \cup \bar{\Omega}_d^+.$$

Now let us have an intuitive analysis on the approximation error for the control. On Ω_d, asymptotically we can assume that

$$0 < B^*p_h + h'(u_h) \to B^*p + h'(u). \tag{5.5}$$

Hence it follows from the optimality conditions that $u = u_h = d$ on Ω_d. Thus the error on Ω_d may be negligible. We should only to estimate the error on

$$\Omega_U \backslash \Omega_d = \Omega_d^- \cup \Omega_d^+$$

in order to avoid over-estimate.

Further it is clear that the states and control approximation errors alone cannot control the approximation errors of numerical coincident sets (see [36] for elliptic obstacle problems). Thus, the measurement of the coincident set approximation errors in our a posteriori error estimates should be considered.

Lemma 5.1. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of optimality conditions (2.6) and semi-discrete DG optimality conditions (3.20), respectively. Assume that $h'(\cdot)$ and $g'(\cdot)$ are locally Lipschitz continuous. Then we have that*

$$|||u - u_h|||^2_{L^2(0,T;L^2(\Omega_U))} \leq C\eta_1^2 + C|||p(u_h) - p_h|||^2_{L^2(L^2)}, \tag{5.6}$$

where

$$\eta_1^2 = \int_0^T \int_{\Omega_d^- \cup \Omega_d^+} (h'(u_h) + B^*p_h)^2,$$

and $y(u_h)$, $p(u_h) \in H^1(0,T; L^2(\Omega)) \cap W$ satisfy $\quad \forall\ w,\ q \in H^1(\Omega)$

$$\begin{cases} (a) \quad (\dfrac{\partial y(u_h)}{\partial t}, w) + A(a(x); y(u_h), w) + J_0^\sigma(y(u_h), w) \\ \qquad\qquad +L(y(u_h),\ w) = (f + Bu_h, w), \\ (b) \quad y(u_h)(x, 0) = y_0^h(x), \qquad y(u_h)|_{\Gamma_-} = 0, \qquad \{a(x)\nabla y(u_h)\} \cdot \nu|_{\Gamma_+} = 0, \\ (c) \quad -(\dfrac{\partial p(u_h)}{\partial t}, q) + A(a(x); q, p(u_h)) + J_0^\sigma(p(u_h), q) \\ \qquad\qquad +L^*(p(u_h),\ q) = (g'(y(u_h)), q), \\ (d) \quad p(u_h)(x, T) = 0, \qquad p(u_h)|_{\Gamma_+} = 0, \qquad \{a(x)\nabla p(u_h)\} \cdot \nu|_{\Gamma_-} = 0. \end{cases} \tag{5.7}$$

Proof. From the uniform convexity of $h(\cdot)$, we have

$$
\begin{aligned}
c|||u-u_h|||^2_{L^2(0,T;L^2(\Omega_U))} &\le \int_0^T (h'(u)-h'(u_h),u-u_h)_U\,dt \\
&= \int_0^T (h'(u)+B^*p,u-u_h)_U\,dt + \int_0^T (h'(u_h)+B^*p(u_h),u_h-u)_U\,dt \\
&\quad + \int_0^T (B^*(p_h-p(u_h)),u-u_h)_U\,dt + \int_0^T (B^*(p(u_h)-p),u-u_h)_U\,dt.
\end{aligned}
\tag{5.8}
$$

Note that the equation (2.6) and (5.7) imply that

$$
\int_0^T (B^*(p(u_h)-p),u-u_h)_U\,dt = \int_0^T \Big(g'(y(u_h))-g'(y),\,y-y(u_h)\Big)\,dt \le 0.
$$

Moreover, note that $u_h\in K^h\subset K$. It follows from (2.6) that

$$
\int_0^T (h'(u)+B^*p,u-u_h)_U\,dt \le 0.
$$

Therefore

$$
\begin{aligned}
c|||u-u_h|||^2_{L^2(0,T;L^2(\Omega_U))} &\le \int_0^T (h'(u_h)+B^*p_h,u_h-u)_U\,dt \\
&\quad + \int_0^T (B^*(p_h-p(u_h)),u-u_h)_U\,dt \\
&:= I_1+I_2.
\end{aligned}
\tag{5.9}
$$

We first estimate I_1. It is clear that for any $t\in(0,T]$

$$
\begin{aligned}
&(h'(u_h)+B^*p_h,u_h-u)_U \\
&= \int_{\Omega_d^-\cup\Omega_d^+} (h'(u_h)+B^*p_h)(u_h-u) + \int_{\Omega_d} (h'(u_h)+B^*p_h)(u_h-u).
\end{aligned}
\tag{5.10}
$$

It is easy to see that

$$
\int_{\Omega_d^-\cup\Omega_d^+} (h'(u_h)+B^*p_h)(u_h-u) \le \frac{1}{2\delta}\int_{\Omega_d^-\cup\Omega_d^+} (h'(u_h)+B^*p_h)^2 + \frac{\delta}{2}|||u_h-u|||^2_{L^2(\Omega_U)},
\tag{5.11}
$$

where δ is an arbitrary small constant.

It follows from the definition of Ω_d that $(h'(d)+B^*p_h)>0$ on Ω_d. Then, we have

$$
\int_{\Omega_d} (h'(u_h)+B^*p_h)(u_h-u) = \int_{\Omega_d} (h'(d)+B^*p_h)(d-u) \le 0.
\tag{5.12}
$$

Thus, (5.10)-(5.12) imply that

$$
I_1 \le C(\delta)\int_0^T\int_{\Omega_d^-\cup\Omega_d^+} (h'(u_h)+B^*p_h)^2 + C\delta|||u_h-u|||^2_{L^2(L^2(\Omega_U))}.
\tag{5.13}
$$

Then for I_2, it is easy to show that

$$
I_2 \le \frac{C}{2\delta}|||p_h-p(u_h)|||^2_{L^2(L^2)} + \frac{\delta}{2}|||u_h-u|||^2_{L^2(L^2(\Omega_U))}.
\tag{5.14}
$$

Taking δ small enough, we obtain from (5.9), (5.13) and (5.14) that

$$|||u-u_h|||^2_{L^2(0,T;L^2(\Omega_U))} \leq C\eta_1^2 + C|||p(u_h)-p_h|||^2_{L^2(L^2)}.$$

This proves (5.6). □

The following lemmas are important in deriving a posteriori error estimates of residual type.

Lemma 5.2. [13] *Let π_h be the standard Lagrange interpolation operator. For $m=0$ or 1, $q>\frac{n}{2}$ and $v\in W^{2,q}(\Omega)$,*

$$|v-\pi_h v|_{W^{2,q}(\Omega^h)} \leq Ch^{2-m}|v|_{W^{2,q}(\Omega^h)}.$$

Lemma 5.3. [37] *For all $v\in W^{1,q}(\Omega^h)$, $1\leq q<\infty$,*

$$|v|_{W^{0,q}(\partial\tau)} \leq C(h_\tau^{-\frac{1}{q}}|v|_{W^{0,q}(\tau)} + h_\tau^{1-\frac{1}{q}}|v|_{W^{1,q}(\tau)}).$$

In order to estimate the error $|||p_h-p(u_h)|||^2_{L^2(0,T;L^2(\Omega))}$ in (5.6), we shall use the following dual equations: For given $F\in L^2(0,T;L^2(\Omega))$,

$$\text{(5.15)} \qquad \begin{cases} \dfrac{\partial\phi}{\partial t} - div(a(x)\nabla\phi) + \beta\cdot\nabla\phi + \alpha\phi = F, & (x,t)\in\Omega\times(0,T], \\ \phi(x,0)=0, \qquad \phi|_{\Gamma_-}=0, \qquad \{a(x)\nabla\phi\}\cdot\nu|_{\Gamma_+}=0, \end{cases}$$

and

$$\text{(5.16)} \qquad \begin{cases} -\dfrac{\partial\psi}{\partial t} - div(a(x)\nabla\psi) - \nabla\cdot(\beta\psi) + \alpha\psi = F, & (x,t)\in\Omega\times(0,T], \\ \psi(x,T)=0, \qquad \psi|_{\Gamma_+}=0, \qquad \{a(x)\nabla\psi\}\cdot\nu|_{\Gamma_-}=0. \end{cases}$$

Under the assumption that $div\beta(x)=0$, $\forall\, x\in\Omega$, we note that for Ω convex, problem (5.16) admits a unique solution $\psi\in H^1(0,T;L^2(\Omega))\cap L^2(0,T;H^2(\Omega))$(see, [38]). There exists the following well known stability results.

Lemma 5.4. [38] *Assume that Ω is a convex domain. Let ϕ and ψ be the solution of (5.15) and (5.16), respectively. Then, for $v=\phi$ or $v=\psi$,*

$$\|v\|_{L^\infty(0,T;L^2(\Omega))} \leq C\|F\|_{L^2(L^2)}, \qquad \|\nabla v\|_{L^2(L^2)} \leq C\|F\|_{L^2(L^2)},$$

$$\|D^2v\|_{L^2(L^2)} \leq C\|F\|_{L^2(L^2)}, \qquad \|\tfrac{\partial v}{\partial t}\|_{L^2(L^2)} \leq C\|F\|_{L^2(L^2)},$$

where $D^2v=\partial^2v/\partial x_i\partial x_j$, $1\leq i,\ j\leq n$.

Now we can provide a proof for $|||p_h-p(u_h)|||^2_{L^2(L^2)}$, similar to that in [40].

Lemma 5.5. *Assume that Ω is a convex domain. Let (y_h,p_h,u_h) be the solutions of (3.19), let $(y(u_h),\,p(u_h))$ be defined by (5.7). Then*

$$\text{(5.17)} \qquad |||y_h-y(u_h)|||^2_{L^2(L^2)} + |||p_h-p(u_h)|||^2_{L^2(L^2)} \leq C\sum_{i=2}^{10}\eta_i^2,$$

where

$$
\begin{cases}
\eta_2^2 & = & \displaystyle\int_0^T \sum_\tau \int_\tau h_\tau^4 \Big(-\frac{\partial p_h}{\partial t} - div(a(x)\nabla p_h) - \nabla\cdot(\beta p_h) + \alpha p_h - g'(y_h)\Big)^2 dx\, dt, \\
\eta_3^2 & = & \displaystyle\int_0^T \sum_{k=1}^{P_h} \int_{e_k} h_{e_k}^3 \{a(x)\nabla p_h \cdot \nu\}^2\, de_k\, dt, \\
\eta_4^2 & = & \displaystyle\int_0^T \sum_{k=1}^{P_h} \int_{e_k} h_{e_k} [p_h]^2\, de_k\, dt, \\
\eta_5^2 & = & \displaystyle\int_0^T \sum_\tau \int_{\partial\tau_+\cap\Gamma_+} h_{e_k}^3 \Big(\big(\beta\cdot n_\tau[p_h]\big)^2 + \big(\beta\cdot n_\tau p_h^+\big)^2\Big)\, de_k\, dt,
\end{cases}
$$

$$
\begin{cases}
\eta_6^2 & = & \displaystyle\int_0^T \sum_\tau \int_\tau h_\tau^4 \Big(\frac{\partial y_h}{\partial t} - div(a(x)\nabla y_h) + \beta\cdot\nabla y_h + \alpha y_h - Bu_h\Big)^2 dx\, dt, \\
\eta_7^2 & = & \displaystyle\int_0^T \sum_{k=1}^{P_h} \int_{e_k} h_{e_k}^3 \{a(x)\nabla y_h \cdot \nu\}^2\, de_k\, dt, \\
\eta_8^2 & = & \displaystyle\int_0^T \sum_{k=1}^{P_h} \int_{e_k} h_{e_k} [y_h]^2\, de_k\, dt, \\
\eta_9^2 & = & \displaystyle\int_0^T \sum_\tau \int_{\partial\tau_-\cap\Gamma_-} h_{e_k}^3 \Big(\big(\beta\cdot n_\tau[y_h]\big)^2 + \big(\beta\cdot n_\tau y_h^+\big)^2\Big)\, de_k\, dt, \\
\eta_{10}^2 & = & \|y_h(x,0) - y_0(x)\|_{L^2}^2,
\end{cases}
$$

where h_{e_k} *is the size of the face* $e_k = \bar\tau_1 \cap \bar\tau_2$*, where* τ_1 *and* τ_2 *are two neighboring elements in* T^h*, and* ν_k *is the unit normal vector on* e_k *outwards* τ_1*.*

Proof. Let ϕ be the solution of (5.15) with $F = p_h - p(u_h)$. Let $\phi_I = \pi_h\phi$ be the interpolation of ϕ defined as in Lemma 5.2. Then it follows from (5.7) and (3.20) that

$$|||p_h - p(u_h)|||_{L^2(0,T;L^2(\Omega))}^2 = \int_0^T (p_h - p(u_h), F)\, dt$$

$$
\begin{aligned}
&= \int_0^T \Big((-\frac{\partial}{\partial t}p_h - div(a(x)p_h) - \nabla\cdot(\beta p_h) + \alpha p_h - g'(y_h),\, \phi - \phi_I)\Big)\, dt \\
&\quad + \int_0^T \Big(\sum_\tau \int_{\partial\tau} (a(x)\nabla p_h\cdot\nu_k)(\phi - \phi_I)\Big)\, dt + \int_0^T \Big(g'(y_h) - g'(y(u_h)),\, \phi\Big)\, dt \qquad (5.18) \\
&\quad - \int_0^T J_0^\sigma(p_h, \phi_I)\, dt + \int_0^T \Big(\sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla\phi_I\cdot\nu_k\}[p_h] - \sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla p_h\cdot\nu_k\}[\phi_I]\Big)\, dt
\end{aligned}
$$

$$+\int_0^T \Big(-\sum_\tau \int_{\partial\tau_+\setminus\Gamma_+} \beta\cdot n_\tau [p_h]\phi_I^+ - \sum_\tau \int_{\partial\tau_+\cap\Gamma_+} \beta\cdot n_\tau p_h^+\phi_I^+\Big)\,dt.$$

$$:= D_1 + D_2 + D_3 + D_4 + D_5 + D_6.$$

Now, we analyze the terms on the right-hand side of (5.18).

(I) It follows from Lemma 5.2 and 5.4 that

$$D_1 \le C(\delta)\eta_2^2 + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}. \tag{5.19}$$

(II) From Lemma 5.2, 5.3 and 5.4, we see that

$$D_2 \le C(\delta)\eta_3^2 + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}, \tag{5.20}$$

where we used

$$\|\phi-\phi_I\|_{L^2(e_k)} \le Ch_{e_k}^{3/2}\|\phi\|_{H^{3/2}(e_k)} \le Ch_{e_k}^{3/2}\|\phi\|_{2,\Omega}.$$

(III) Lemma 5.4 and Schwartz inequality imply that

$$D_3 \le C(\delta)|||y_h - y(u_h)|||^2_{L^2(L^2)} + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}. \tag{5.21}$$

(IV) Similarly like D_2, we derive

$$D_4 \le C(\delta)\eta_4^2 + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}. \tag{5.22}$$

(V) Similarly, we get

$$\begin{aligned} D_5 &= |-\int_0^T \Big(\sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla(\phi-\phi_I)\cdot\nu_k\}[p_h] + \sum_{k=1}^{P_h}\int_{e_k}\{a(x)\nabla p_h\cdot\nu_k\}[\phi-\phi_I]\Big)\,dt| \\ &\le C(\delta)\eta_3^2 + C(\delta)\eta_4^2 + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}. \end{aligned} \tag{5.23}$$

(VI) Also, we have

$$\begin{aligned} D_6 &= |\int_0^T \Big(\sum_\tau \int_{\partial\tau_+\setminus\Gamma_+} \beta\cdot n_\tau [p_h](\phi-\phi_I^+) + \sum_\tau \int_{\partial\tau_+\cap\Gamma_+} \beta\cdot n_\tau p_h^+(\phi-\phi_I^+)\Big)\,dt| \\ &\le C(\delta)\eta_5^2 + C\delta|||p_h - p(u_h)|||^2_{L^2(L^2)}. \end{aligned} \tag{5.24}$$

Then letting δ be small enough, it follows from (5.19)-(5.24) that

$$|||p_h - p(u_h)|||^2_{L^2(L^2)} \le C\sum_{i=2}^{5}\eta_i^2 + C|||y_h - y(u_h)|||^2_{L^2(L^2)}. \tag{5.25}$$

Let ψ be the solution of (5.16) with $F = y_h - y(u_h)$. Similarly to the proof of (5.25), we have that

$$|||y_h - y(u_h)|||^2_{L^2(L^2)} = \int_0^T (y_h - y(u_h), F)\,dt$$

$$\leq C(\delta)\int_0^T \sum_\tau \int_\tau h_\tau^4\Big(\frac{\partial y_h}{\partial t} - div(a(x)\nabla y_h) + \beta\cdot\nabla y_h + \alpha y_h - Bu_h\Big)^2 dx\,dt,$$

$$+C(\delta)\int_0^T \sum_{k=1}^{P_h}\int_{e_k} h_{e_k}^3\{a(x)\nabla y_h\cdot\nu\}^2\,de_k\,dt + C(\delta)\int_0^T\sum_{k=1}^{P_h}\int_{e_k} h_{e_k}[y_h]^2\,de_k\,dt$$

$$+C(\delta)\int_0^T\sum_\tau\int_{\partial\tau_-\cap\Gamma_-} h_{e_k}^3\Big(\big(\beta\cdot n_\tau[y_h]\big)^2 + \big(\beta\cdot n_\tau y_h^+\big)^2\Big)\,de_k\,dt + C\delta\|\psi(x,0)\|^2$$

$$+C(\delta)\|y_h(x,0) - y_0(x)\|^2_{L^2} + C\delta\int_0^T \|\psi\|^2_{2,\Omega}.$$

Hence, letting δ be small enough, we have

$$|||y_h - y(u_h)|||^2_{L^2(0,T;L^2(\Omega))} \leq C\sum_{i=6}^{10}\eta_i^2. \tag{5.26}$$

Then, (5.17) follows from (5.25) and (5.26). □

From Lemma 5.1 and Lemma 5.4, we have the following a posteriori error estimates.

Theorem 5.1. *Let (y,p,u) and (y_h,p_h,u_h) be the solutions of optimality conditions (2.6) and semi-discrete DG optimality conditions (3.20), respectively. Assume that all the conditions in Lemma 5.1 and 5.5 are valid. Then*

$$|||y_h - y|||^2_{L^2(L^2)} + |||p_h - p|||^2_{L^2(L^2)} + |||u_h - u|||^2_{L^2(L^2(\Omega_U))} \leq C\sum_{i=2}^{10}\eta_i^2, \tag{5.27}$$

where η_1 is defined in Lemma 5.1, η_i, $i = 2,\cdots,10$, are defined in Lemma 5.5.

Proof. By (5.6),(5.25) and (5.26), we can get

$$|||u - u_h|||^2_{L^2(L^2(\Omega_U))} \leq C\eta_1^2 + C|||p(u_h) - p_h|||^2_{L^2(L^2)} \leq C\sum_{i=1}^{10}\eta_i^2. \tag{5.28}$$

Note that

$$\begin{aligned}|||y - y_h|||_{L^2(L^2)} &\leq |||y - y(u_h)|||_{L^2(L^2)} + |||y(u_h) - y_h|||_{L^2(L^2)},\\ |||p - p_h|||_{L^2(L^2)} &\leq |||p - p(u_h)|||_{L^2(L^2)} + |||p(u_h) - p_h|||_{L^2(L^2)},\end{aligned}$$

and

$$|||y - y_h|||_{L^2(L^2)} \leq C|||u - u_h|||_{L^2(L^2(\Omega_U))},$$

$$|||p - p_h|||_{L^2(L^2)} \leq C|||y - y(u_h)|||_{L^2(L^2)} \leq C|||u - u_h|||_{L^2(L^2(\Omega_U))}.$$

Then, it follows from Lemma 5.5 and (5.28) that

$$|||y - y_h|||_{L^2(L^2)} + |||p - p_h|||_{L^2(L^2)} \leq C\sum_{i=1}^{10}\eta_i^2. \tag{5.29}$$

Thus, (5.27) follows (5.28) and (5.29). □

Acknowledgments: The author expresses thanks to Prof. W.B. Liu for fruitful discussion and suggestions in preparing this manuscript. The author also expresses thanks to the referees for their helpful suggestions, which lead to improvements of the presentation.

References

[1] W.Alt and U. Mackenroth, Convergence of finite element approximations to state constrained convex parabolic boundary control problem, *SIAM J. Control Optim.,* 27(1989), 718-736.

[2] F.S. Falk, Approximation of a class of optimal control problems with order of convergence estimates, *J. Math. Anal. Appl.,* 44(1973), 28-47.

[3] D.A. French and J.T.King, Approximation of an elliptic control problem by the finite element method, *Numer.Funct. Anal. Optim.,* 12(1991), 299-314.

[4] T. Geneci, On the approximation of the solution of an optimal control problem governed by an elliptic equation, *RAIRO Anal. Numer.,* 13(1979), 313-328.

[5] G. Knowles, Finite element approximation of parabolic time optimal control problems, *SIAM J. Control Optim.,* 20(1982), 414-427.

[6] M.D. Gunzburger and L.S. Hou, Finite-dimensional approximation of a class of constrained nonliear control problems, *SIAM J. Control Optim.,* 34(1996), 1001-1043.

[7] N. Arada, E. Casas, and F. Troltzsch, Error estimates for a semilinear elliptic control problem, *Compt. Optim. Appl.,* 23(2002), 201-229.

[8] W.B. Liu, and D., Ti"a, Error estimates for the finite element approximation of a class of nonlinear optimal control problems, *J. Numer. Func. Optim.,* 22(2001), 953-972.

[9] W.B. Liu, and N. N. Yan, A posteriori error estimators for a class of variational inequalities, *J. Sci. Comp.,* 35(2000), 361-393.

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

[11] W.B. Liu, and N. N. Yan, A posteriori error estimates for control problems governed by Stokes' equation, *SIAM J. Numer. Anal.,* 40(2003), 1805-1869.

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

[13] P.G. Ciarlet, The finite element method for elliptic problems, North-Holland Publ., Amsterdam,1978.

[14] J. Haslinger and P. Neittanmaki, Finite element approximation for optimal shape design, John Wiley and Sons, Chichester, UK, 1989.

[15] P. Neittanmaki and D. Tiba, Optimal control of nonlinear parabolic systems: Theory, Algorithms and Applications, M. Dekker, New York, 1994.

[16] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems, *Comput. Meth. App. Mech. Engrg.,* 107 (1993), 117-129.

[17] B. Rivière, M.F.Wheeler, V.Girault, A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems, *SIAM J. Numer. Anal.,* 39(3)(2001), 902-931.

[18] C. Dawson and J. Proft, Discontinuous and coupled continuous/discontinuous Galerkin methods for the shallow water equations, *Comput. Meth. Appl. Mech. Engrg.,* 191(2002), 4721-4746.

[19] G.Baker, Finite element methods for elliptic equations using nonconforming elements, *Math. Comp.,* 31(1977), 45-59.

[20] J.Douglas,Jr. and T.Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, Lecture Notes in Physics, 58(1976), 207-216.

[21] M.F.Wheeler , An elliptic collocation-finite element method with interior penalties, *SIAM J. Numer. Anal.,* 15(1978), 152-161.

[22] D.N.Arnold, An interior penalty finite element method with discontinuous elements, *SIAM J. Numer. Anal.,* 19(1982), 724-760.

[23] J.A. Nitsche, *Ü*ber ein Variationspringzip zur Lösung von Drichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, Abh. Math. Sem. Univ. Hamburg. 36(1971), 9-15.

[24] J.Douglas,Jr., M.F.Wheeler, B.L. Darlow, R.P.Kendall, Self-adaptive finite element simulation of miscible displacement in porous media, *Comput. Meth. Appl. Mech. Engrg.,* 47(1984), 131-159.

[25] B.Rivièrinterior penalty, constrained and discontinuous Galerkin methods for elliptic problems, *Computational Geosciences,* 8(1999), 337-360.

[26] B.Rivière, M.F.Wheeler, K.Banas, Part II. Discontinuous Galerkin method applied to single phase flow in porous media, *Computational Geosciences,* 4(4)(2000), 337-341.

[27] S.S. Collis, M. Heinkenschloss, Analysis of the streamline upwind/Petrov Galerkin method applied to the solution of optimal control problems. CAAM TR02-01 (2002).

[28] L. Dedè, A. Quarteroni, Optimal control and numerical adaptivity for advection-diffusion equations. *M2AN,* 39(2005), 1019-1040.

[29] A. Quarteroni, G. Rozza, L. Ded and A. Quaini, Numerical approximation of a control problem for advection-diffusion processes. System Modeling and Optimization, 261-273, IFIP International Federation for Information Processing, 199, Springer, New York, 2006.

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

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

[32] J.L. Lions, E. Magenes, Non homogeneous boundary value problems and applications, Grandlehre B. 181, Springer-Verlag, 1972.

[33] R.A.Adams, Sobolev spaces, Academic Press, New York, 1975.

[34] P.Houston, CH. Schwab, and E. Süli, Stabilized *hp*-finite element methods for first-order hyperbolic problems, *SIAM J. Numer. Anal.,* 37(5)(2000), 1618-1643.

[35] P.Houston, CH. Schwab, and E. Süli, Discontinuous *hp*-finite element methods for advection-diffusion-reaction, *SIAM J. Numer. Anal.,* 39(6)(2002), 2133-2163.

[36] A. Veeser, Efficient and reliable a posteriori error estimates for elliptic obstacle problems, *SIAM J., Numer. Anal.,* 39(6)(2001), 146-167.

[37] A. Kufner, O.John, and S. Fucik, Function spaces. Nordhoff, Leyden, The Netherlands, 1977.

[38] P.Houston,and E. Süli, A posteriori error analysis for linear convection-diffusion problems under weak mesh regularity assumptions. Report 97/03, Oxford University Computing Laboratory.

[39] P. Houston. Lagrange-Galerkin methods for unsteady convection-diffusion problems: a posteriori error analysis and adaptivity. PhD thesis, University of Oxford, 1996.

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

School of Mathematics, Shandong University, Jinan 250100, P.R.China
E-mail: `tjsun@sdu.edu.cn`