12-1-2016

Error Analysis of an HDG Method for a Distributed Optimal

Huiqing Zhu
University of Southern Mississippi, Huiqing.Zhu@usm.edu

Fatih Celiker
Wayne State University, celiker@math.wayne.edu

ERROR ANALYSIS OF AN HDG METHOD FOR A DISTRIBUTED OPTIMAL CONTROL PROBLEM

HUIQING ZHU AND FATIH CELIKER

Abstract. In this paper, we present *a priori* error analysis of a hybridizable discontinuous Galerkin (HDG) method for a distributed optimal control problem governed by diffusion equations. The error estimates are established based on the projection-based approach recently used to analyze these methods for the diffusion equation. We proved that for approximations of degree k on conforming meshes, the orders of convergence of the approximation to fluxes and scalar variables are $k+1$ when the local stabilization parameter is suitably chosen.

1. Introduction

Optimal control problems governed by partial differential equations arise in many scientific and engineering computing problems such as aerodynamics [10, 23], medicine [1, 14], and mathematical finance [2, 9], to name but a few. The mathematical foundations for problems of this type were set down by J.L. Lions in the 1960s [16]. In the past few decades, there has been a considerable amount of work concentrating on numerical solutions of optimal control problems [8, 21, 24]. Among different numerical methods, finite element approximation of optimal control problems have been extensively studied. Some *a priori* and *a posteriori* error analysis can be found in [5, 11, 12, 19, 24] and references cited therein. Recently, discontinuous Galerkin methods have also been applied for a few optimal control problems [15, 20, 25, 26].

Hybridizable discontinuous Galerkin (HDG) methods were proposed by Cockburn *et al.* in [6] as an improvement of traditional discontinuous Galerkin methods. The main advantage of these methods is that the only globally coupled degrees of freedom are the ones on the element boundaries, which substantially reduces the computational cost. HDG methods also produce optimal approximations not only to the potential but also to the flux for elliptic problems [7]. Furthermore, HDG methods have many other desirable properties such as ability to handle complex geometries and high order approximations, stability and low dispersion for discretizations of hyperbolic systems, simple imposition of boundary conditions. They also have superconvergence properties, which in turn delivers efficient postprocessing techniques.

In this paper, instead of aiming for maximal generality, we will concentrate on a specific HDG method for optimal control problems governed by a model elliptic partial differential equation. We consider this as a stepping stone towards devising

1991 *Mathematics Subject Classification.* 49K20, 49M25, 65J10, 65N15.

Key words and phrases. hybridizable discontinuous Galerkin method, optimal control, error estimates.

Corresponding author: Huiqing Zhu, Huiqing.Zhu@usm.edu.

HDG methods for more complicated optimal control problems. We also intend to use this for exploring potential advantages of HDG methods when applied to problems of this class. Our major motivation for applying HDG methods to these problems is the established fact that they have superior stability and convergence properties for many classical partial differential equations such as convection-diffusion-reaction problems, elasticity problems, Stokes and Navier-Stokes equations, to name a few. Since many optimal control problems are governed by similar partial differential equations, our expectation is that desirable properties of HDG methods will also carry over to these optimal control problems. The main result of this paper actually proves, at least for this simple case, that this hope is not in vain, indicating that the same will potentially hold for more complicated problems.

Next, we describe the optimal control problem for which we will devise and analyze an HDG method. Let Ω be a Lipschitz polyhedral domain in $\mathbb{R}^n$ with $n \geq 2$. Given $f, \widetilde{y} \in L^2(\Omega)$ and $g \in H^{3/2}(\partial\Omega)$ we let

$$J(y^*, u^*) := \frac{1}{2}\|y^* - \widetilde{y}\,\|^2 + \frac{\alpha}{2}\|u^*\|^2$$

and

$$(y, u) = \arg\min J(y^*, u^*) \tag{1.1}$$

for all $(y^*, u^*) \in Y \times U$ subject to

$$\begin{aligned} -\nabla \cdot (a\nabla y^*) &= f + u^* && \text{in } \Omega, \\ y^* &= g && \text{on } \partial\Omega, \end{aligned} \tag{1.2}$$

where $Y := \{w \in H^1(\Omega) \mid w = g \text{ on } \partial\Omega\}$, $U := L^2(\Omega)$. Furthermore, $a > 0$ and $\alpha > 0$ are given diffusion and regularization parameters, respectively. We denote by $\|\cdot\|$ the usual L^2 norm on Ω.

To define the HDG approximation of the optimal control problem (1.1)–(1.2), we need a weak formulation for the state equation (1.2). We denote the L^2-inner products on $L^2(D)$ and $L^2(\partial D)$ by $(v, w)_D$ and $\langle v, w\rangle_{\partial D}$, respectively, where D is an arbitrary subdomain of Ω. We will drop the subscript if $D = \Omega$. With this notation the standard weak formulation for the state equation reads as follows: given $f \in L^2(\Omega)$, find $y^*(u^*) \in Y$ such that

$$(a\nabla y^*, \nabla w) = (f + u^*, w), \quad \forall\, w \in W := H_0^1(\Omega). \tag{1.3}$$

It is well known that the theory in [17](Sec. II.1) guarantees the existence of a unique solution $(y, u) \in H^1(\Omega) \times L^2(\Omega)$ of (1.1) and (1.3).

The state y and the control u solve the optimal control problem (1.1) and (1.3) if and only if there exists an adjoint $z \in H_0^1(\Omega)$ such that y, u, z satisfy the state equation

$$(a\nabla y, \nabla w_1) = (f + u, w_1), \quad \forall\, w_1 \in W, \tag{1.4a}$$

the adjoint equation,

$$(a\nabla z, \nabla w_2) = (\widetilde{y} - y, w_2), \quad \forall\, w_2 \in W, \tag{1.4b}$$

and the gradient equation,

$$(\alpha u - z, r) = 0, \qquad \forall\, r \in L^2(\Omega). \tag{1.4c}$$

Since (1.4c) is merely an algebraic equation and $W \subset L^2(\Omega)$, it is possible to eliminate u from these equations by setting $u = \beta z$ where $\beta := \alpha^{-1}$ and rewriting (1.4a) and (1.4b) in the equivalent form

$$(a\nabla y, \nabla w_1) = (f + \beta z, w_1), \quad \forall\, w_1 \in W, \tag{1.5a}$$

$$(a\nabla z, \nabla w_2) = (\widetilde{y} - y, w_2), \quad \forall\, w_2 \in W. \tag{1.5b}$$

The main goal of this paper is to devise and prove a priori error estimates for an HDG method for (1.5).

The rest of the paper is organized as follows: In Sec. 2, we introduce the HDG formulation. In Sec. 3, we present the error estimates of the HDG discretization, namely, the main result of this paper. A detailed proof of the main result will be given in Sec. 4. We end by some concluding remarks in Sec. 5.

2. HDG formulation

Let Ω_h be a regular partitioning of Ω and K be an element in Ω_h. We denote the diameter of K by h_K and set $h = \max_{K\in\Omega_h} h_K$. We further denote by $\mathcal{E}_h$ and $\mathcal{E}_h^\circ$ boundaries and interior boundaries of Ω_h, respectively. We will work in the following finite element spaces

$$\begin{aligned}
\boldsymbol{V}_h &:= \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega_h) \mid \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}_k(K),\ \forall\, K \in \Omega_h\},\\
W_h &:= \{w \in L^2(\Omega_h) \mid w|_K \in \mathcal{P}_k(K),\ \forall\, K \in \Omega_h\},\\
M_h &:= \{\mu \in L^2(\mathcal{E}_h) \mid \mu|_e \in \mathcal{P}_k(e),\ \forall\, e \in \mathcal{E}_h^\circ;\ \mu|_e = 0,\ \forall\, e \in \mathcal{E}_h \cap \partial\Omega\}.
\end{aligned}$$

Here, $\mathcal{P}_k(K)$ and $\mathcal{P}_k(e)$ are spaces of polynomial of total degree at most $k \geq 0$ on K and on e, respectively. The space of vector valued polynomial functions is $\boldsymbol{\mathcal{P}}_k(K) := [\mathcal{P}_k(K)]^n$.

An intermediate step for defining the HDG approximation of (1.5) is writing the strong form of the system (1.5) as the following first-order system of differential equations. The state equation and the gradient equation

$$c\boldsymbol{p} + \nabla y = 0 \ \text{ in } \Omega, \tag{2.1a}$$

$$\nabla\cdot\boldsymbol{p} - \beta z = f \ \text{ in } \Omega, \tag{2.1b}$$

$$y = g \ \text{ on } \partial\Omega, \tag{2.1c}$$

and the adjoint equation

$$c\boldsymbol{r} + \nabla z = 0 \ \text{ in } \Omega, \tag{2.1d}$$

$$\nabla\cdot\boldsymbol{r} + y = \widetilde{y} \ \text{ in } \Omega, \tag{2.1e}$$

$$z = 0 \ \text{ on } \ \partial\Omega, \tag{2.1f}$$

where $c := a^{-1}$. Notice that we have introduced two more unknowns, $\boldsymbol{p}$ and $\boldsymbol{r}$, into the system. Furthermore, we will introduce two more in the following step. This proliferation of unknowns, however, will be greatly compensated for through the hybridization process in which we eliminate all the *internal* degrees of freedom.

Remark 2.1. Note that the strong form of (1.4c), namely,

$$\alpha u - z = 0 \quad \text{in } \Omega \tag{2.2}$$

is imbedded in (2.1b). We will thus resort to a slight abuse of notation as we mention u as part of the solution of the governing equation (2.1) since once z is obtained, one can easily recover u from (2.2).

The HDG method seeks an approximation $(\boldsymbol{p}_h, y_h, \widehat{y}_h, \boldsymbol{r}_h, z_h, \widehat{z}_h)$ to the exact solution $(\boldsymbol{p}, y, y|_{\mathcal{E}_h}, \boldsymbol{r}, z, z|_{\mathcal{E}_h})$ of (2.1) in the space $\boldsymbol{V}_h \times W_h \times M_h \times \boldsymbol{V}_h \times W_h \times M_h$ such that

$$
\begin{aligned}
(c\boldsymbol{p}_h, \boldsymbol{v}_1) \ - (y_h, \nabla\cdot\boldsymbol{v}_1) + \langle \widehat{y}_h, \boldsymbol{v}_1\cdot\boldsymbol{n}\rangle \ &= 0, && (2.3a)\\
-(\boldsymbol{p}_h, \nabla w_1) - (\beta z_h, w_1) \quad + \langle \widehat{\boldsymbol{p}}_h\cdot\boldsymbol{n}, w_1\rangle &= (f, w_1), && (2.3b)
\end{aligned}
$$

for all $(\boldsymbol{v}_1, w_1)$ in $\boldsymbol{V}_h \times W_h$

$$
\begin{aligned}
(c\,\boldsymbol{r}_h, \boldsymbol{v}_2) \ - (z_h, \nabla\cdot\boldsymbol{v}_2) + \langle \widehat{z}_h, \boldsymbol{v}_2\cdot\boldsymbol{n}\rangle \ &= 0, && (2.3c)\\
-(\boldsymbol{r}_h, \nabla w_2) + (y_h, w_2) \qquad + \langle \widehat{\boldsymbol{r}}_h\cdot\boldsymbol{n}, w_2\rangle &= (\widetilde{y}, w_2), && (2.3d)
\end{aligned}
$$

for all $(\boldsymbol{v}_2, w_2)$ in $\boldsymbol{V}_h \times W_h$, and

$$
\begin{aligned}
\langle \widehat{\boldsymbol{p}}_h\cdot\boldsymbol{n}, \mu_1\rangle \qquad &= 0, && (2.3e)\\
\langle \widehat{\boldsymbol{r}}_h\cdot\boldsymbol{n}, \mu_2\rangle \qquad &= 0, && (2.3f)
\end{aligned}
$$

for all (μ_1, μ_2) in $M_h \times M_h$. Recall that, for a vector-valued function $\boldsymbol{v}$ and a scalar function w defined on Ω_h

$$
\langle \boldsymbol{v}\cdot\boldsymbol{n}, w\rangle = \langle \boldsymbol{v}\cdot\boldsymbol{n}, w\rangle_{\partial\Omega_h} = \sum_{K\in\Omega_h} \langle \boldsymbol{v}\cdot\boldsymbol{n}, w\rangle_{\partial K}
$$

where $\boldsymbol{n}$ appearing in the boundary integrals inside the summation denotes the unit outward normal vector to the boundary of the element K. The *numerical traces* on $\partial\Omega_h$ are defined as

$$
\begin{aligned}
\widehat{\boldsymbol{p}}_h &= \boldsymbol{p}_h + \sigma(y_h - \widehat{y}_h)\boldsymbol{n},\\
\widehat{\boldsymbol{r}}_h &= \boldsymbol{r}_h + \sigma(z_h - \widehat{z}_h)\boldsymbol{n},
\end{aligned} \tag{2.4}
$$

where σ is a nonnegative stabilization function defined on $\partial\Omega_h$, which we assume to be constant on each face of the triangulation. Observe that $\widehat{\boldsymbol{p}}_h$ and $\widehat{\boldsymbol{r}}_h$ are possibly double-valued on $\mathcal{E}_h^\circ$. For example, when evaluating (2.3e),

$$
\langle \widehat{\boldsymbol{p}}_h\cdot\boldsymbol{n}, \mu_1\rangle = \sum_{K\in\Omega_h} \langle \widehat{\boldsymbol{p}}_h\cdot\boldsymbol{n}, \mu_1\rangle_{\partial K} = \sum_{K\in\Omega_h} \langle \boldsymbol{p}_h\cdot\boldsymbol{n} + \sigma(y_h - \widehat{y}_h), \mu_1\rangle_{\partial K}
$$

the values of $\boldsymbol{p}_h$ and y_h from within the element K are used inside the summation, and $\boldsymbol{n}$ is the unit outward normal vector to ∂K as above. Note, however, that $\widehat{y}_h$ (as well as $\widehat{z}_h$) is single-valued on $\mathcal{E}_h$.

Denoting the L^2-orthogonal projection onto M_h by P_M, the boundary condition (2.1c) is enforced *weakly* by requiring that

$$
\widehat{y}_h = P_M g \quad \text{on } \partial\Omega,
$$

and the boundary condition (2.1f) is similarly enforced by requiring that

$$
\widehat{z}_h = 0 \quad \text{on } \partial\Omega.
$$

This completes the definition of the HDG methods that we will consider in this paper.

Remark 2.2. In the spirit of Remark 2.1, the HDG formulation (2.3) also defines an approximation $u_h \in W_h$ to u such that

$$\alpha u_h - z_h = 0 \quad \text{in } \Omega_h,$$

or in its weak from

$$(\alpha u_h - z_h, w) = 0 \quad \forall\, w \in W_h. \tag{2.5}$$

We will thus consider (2.5) as part of the HDG formulation (2.3). Notwithstanding the fact that when implementing these methods one would not include (2.5) as part of the system of equations but rather recover u_h from z_h, we will state and prove error estimates on u_h under the premises detailed above.

The formulation (2.3) together with (2.4) is sufficient for the error analysis that will be carried out in Sec. 4. However, we would like to elucidate on a point that has been mentioned earlier, namely, efficient implementation of these methods. To this end, we have to define four *local solvers*. The reason why they are called local is that they are defined on a single element K in Ω_h and hence their computational cost is very low and it is extremely parallelizable. The first local solver is the mapping

$$\widehat{y} \mapsto (\boldsymbol{P}_{\widehat{y}}, Y_{\widehat{y}}, \boldsymbol{R}_{\widehat{y}}, Z_{\widehat{y}}) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K) \times \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$$

such that

$$\begin{aligned}
(c\boldsymbol{P}_{\widehat{y}}, \boldsymbol{v}_1)_K \quad - (Y_{\widehat{y}}, \nabla\cdot \boldsymbol{v}_1)_K + \langle \widehat{y}, \boldsymbol{v}_1\cdot \boldsymbol{n}\rangle_{\partial K} &= 0, && \text{(2.6a)}\\
-(\boldsymbol{P}_{\widehat{y}}, \nabla w_1)_K - (\beta Z_{\widehat{y}}, w_1)_K \quad + \langle \widehat{\boldsymbol{P}}_{\widehat{y}}\cdot \boldsymbol{n}, w_1\rangle_{\partial K} &= 0, && \text{(2.6b)}\\
(c\,\boldsymbol{R}_{\widehat{y}}, \boldsymbol{v}_2)_K \quad - (Z_{\widehat{y}}, \nabla\cdot \boldsymbol{v}_2)_K &= 0, && \text{(2.6c)}\\
-(\boldsymbol{R}_{\widehat{y}}, \nabla w_2)_K + (Y_{\widehat{y}}, w_2)_K \quad + \langle \widehat{\boldsymbol{R}}_{\widehat{y}}\cdot \boldsymbol{n}, w_2\rangle_{\partial K} &= 0, && \text{(2.6d)}
\end{aligned}$$

for all $(\boldsymbol{v}_1, w_1)$, $(\boldsymbol{v}_2, w_2) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$. Here,

$$\begin{aligned}
\widehat{\boldsymbol{P}}_{\widehat{y}} &= \boldsymbol{P}_{\widehat{y}} + \sigma(Y_{\widehat{y}} - \widehat{y})\boldsymbol{n},\\
\widehat{\boldsymbol{R}}_{\widehat{y}} &= \boldsymbol{R}_{\widehat{y}} + \sigma(Z_{\widehat{y}})\boldsymbol{n}.
\end{aligned} \tag{2.7}$$

Note that (2.6) is nothing but the restriction to the element K of (2.3a)–(2.3d) with $\widehat{y}_h = \widehat{y}$, $\widehat{z}_h = 0$, $f = 0$, and $\widetilde{y} = 0$. The definition of the numerical traces (2.7) is also in agreement with (2.4). In this sense, this local solver picks up information that is relevant only to $\widehat{y}_h$. Analogously, the second local solver is be designed to pick up information relevant to $\widehat{z}_h$. Thus, it is the mapping

$$\widehat{z} \mapsto (\boldsymbol{P}_{\widehat{z}}, Y_{\widehat{z}}, \boldsymbol{R}_{\widehat{z}}, Z_{\widehat{z}}) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K) \times \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$$

such that

$$\begin{aligned}
(c\boldsymbol{P}_{\widehat{z}}, \boldsymbol{v}_1)_K \quad - (Y_{\widehat{z}}, \nabla\cdot \boldsymbol{v}_1)_K &= 0, && \text{(2.8a)}\\
-(\boldsymbol{P}_{\widehat{z}}, \nabla w_1)_K - (\beta Z_{\widehat{z}}, w_1)_K \quad + \langle \widehat{\boldsymbol{P}}_{\widehat{z}}\cdot \boldsymbol{n}, w_1\rangle_{\partial K} &= 0, && \text{(2.8b)}\\
(c\,\boldsymbol{R}_{\widehat{z}}, \boldsymbol{v}_2)_K \quad - (Z_{\widehat{z}}, \nabla\cdot \boldsymbol{v}_2)_K + \langle \widehat{z}, \boldsymbol{v}_1\cdot \boldsymbol{n}\rangle_{\partial K} &= 0, && \text{(2.8c)}\\
-(\boldsymbol{R}_{\widehat{z}}, \nabla w_2)_K + (Y_{\widehat{z}}, w_2)_K \quad + \langle \widehat{\boldsymbol{R}}_{\widehat{z}}\cdot \boldsymbol{n}, w_2\rangle_{\partial K} &= 0, && \text{(2.8d)}
\end{aligned}$$

for all $(\boldsymbol{v}_1, w_1)$, $(\boldsymbol{v}_2, w_2) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$. Here,

$$\begin{aligned}
\widehat{\boldsymbol{P}}_{\widehat{z}} &= \boldsymbol{P}_{\widehat{z}} + \sigma(Y_{\widehat{z}})\boldsymbol{n},\\
\widehat{\boldsymbol{R}}_{\widehat{z}} &= \boldsymbol{R}_{\widehat{z}} + \sigma(Z_{\widehat{z}} - \widehat{z})\boldsymbol{n}.
\end{aligned} \tag{2.9}$$

The remaining two local solvers are also defined in the same spirit, that is, to pick up information relevant to f and $\widetilde{y}$, respectively. The third one is the mapping

$$f \mapsto (\boldsymbol{P}_f, Y_f, \boldsymbol{R}_f, Z_f) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K) \times \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$$

such that

$$(c\boldsymbol{P}_f, \boldsymbol{v}_1)_K \quad - (Y_f, \nabla \cdot \boldsymbol{v}_1)_K \qquad = 0, \tag{2.10a}$$

$$-(\boldsymbol{P}_f, \nabla w_1)_K - (\beta Z_f, w_1)_K \quad + \langle \widehat{\boldsymbol{P}}_f \cdot \boldsymbol{n}, w_1 \rangle_{\partial K} = f, \tag{2.10b}$$

$$(c\,\boldsymbol{R}_f, \boldsymbol{v}_2)_K \quad - (Z_f, \nabla \cdot \boldsymbol{v}_2)_K \qquad = 0, \tag{2.10c}$$

$$-(\boldsymbol{R}_f, \nabla w_2)_K + (Y_f, w_2)_K \quad + \langle \widehat{\boldsymbol{R}}_f \cdot \boldsymbol{n}, w_2 \rangle_{\partial K} = 0, \tag{2.10d}$$

for all $(\boldsymbol{v}_1, w_1)$, $(\boldsymbol{v}_2, w_2) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$. Here,

$$\begin{aligned} \widehat{\boldsymbol{P}}_f &= \boldsymbol{P}_f + \sigma(Y_f)\boldsymbol{n}, \\ \widehat{\boldsymbol{R}}_f &= \boldsymbol{R}_f + \sigma(Z_f)\boldsymbol{n}. \end{aligned} \tag{2.11}$$

The fourth local solver is the mapping

$$\widetilde{y} \mapsto (\boldsymbol{P}_{\widetilde{y}}, Y_{\widetilde{y}}, \boldsymbol{R}_{\widetilde{y}}, Z_{\widetilde{y}}) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K) \times \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$$

such that

$$(c\boldsymbol{P}_{\widetilde{y}}, \boldsymbol{v}_1)_K \quad - (Y_{\widetilde{y}}, \nabla \cdot \boldsymbol{v}_1)_K \qquad = 0, \tag{2.12a}$$

$$-(\boldsymbol{P}_{\widetilde{y}}, \nabla w_1)_K - (\beta Z_{\widetilde{y}}, w_1)_K \quad + \langle \widehat{\boldsymbol{P}}_{\widetilde{y}} \cdot \boldsymbol{n}, w_1 \rangle_{\partial K} = 0, \tag{2.12b}$$

$$(c\,\boldsymbol{R}_{\widetilde{y}}, \boldsymbol{v}_2)_K \quad - (Z_{\widetilde{y}}, \nabla \cdot \boldsymbol{v}_2)_K \qquad = 0, \tag{2.12c}$$

$$-(\boldsymbol{R}_{\widetilde{y}}, \nabla w_2)_K + (Y_{\widetilde{y}}, w_2)_K \quad + \langle \widehat{\boldsymbol{R}}_{\widetilde{y}} \cdot \boldsymbol{n}, w_2 \rangle_{\partial K} = \widetilde{y}, \tag{2.12d}$$

for all $(\boldsymbol{v}_1, w_1)$, $(\boldsymbol{v}_2, w_2) \in \boldsymbol{\mathcal{P}}_k(K) \times \mathcal{P}_k(K)$. Here,

$$\begin{aligned} \widehat{\boldsymbol{P}}_{\widetilde{y}} &= \boldsymbol{P}_{\widetilde{y}} + \sigma(Y_{\widetilde{y}})\boldsymbol{n}, \\ \widehat{\boldsymbol{R}}_{\widetilde{y}} &= \boldsymbol{R}_{\widetilde{y}} + \sigma(Z_{\widetilde{y}})\boldsymbol{n}. \end{aligned} \tag{2.13}$$

The hybridization process is then as follows. We first solve (2.3e)–(2.3f) for the unknowns $\widehat{y}_h$ and $\widehat{z}_h$. Note that this system of equations is the only global system of equations that needs to be solved and it involves degrees of freedom only on element faces and is independent of all internal degrees of freedom[1]. We then recover the remaining unknowns (internal degrees of freedom) in a postprocessing step by setting

$$\boldsymbol{p}_h = \boldsymbol{P}_{\widehat{y}_h} + \boldsymbol{P}_{\widehat{z}_h} + \boldsymbol{P}_f + \boldsymbol{P}_{\widetilde{y}}, \tag{2.14a}$$

$$y_h = Y_{\widehat{y}_h} + Y_{\widehat{z}_h} + Y_f + Y_{\widetilde{y}}, \tag{2.14b}$$

$$\boldsymbol{r}_h = \boldsymbol{R}_{\widehat{y}_h} + \boldsymbol{R}_{\widehat{z}_h} + \boldsymbol{R}_f + \boldsymbol{R}_{\widetilde{y}}, \tag{2.14c}$$

$$z_h = Z_{\widehat{y}_h} + Z_{\widehat{z}_h} + Z_f + Z_{\widetilde{y}}. \tag{2.14d}$$

Observe that this last step can be achieved in an element-by-element fashion and hence its computational cost is negligible.

[1]This is due to the fact that, by (2.14), all internal degrees of freedom appearing in (2.3e) and (2.3f) can be expressed in terms of the given data (f and $\widetilde{y}$) and the unknowns $\widehat{y}_h$ and $\widehat{z}_h$.

3. The Main Result

We begin with introducing the projection operator

$$\Pi_h(\boldsymbol{q},\psi) := (\boldsymbol{\Pi}_V\boldsymbol{q}, \Pi_W\psi)$$

defined in [7] which will be instrumental in our proof. Here $\boldsymbol{\Pi}_V\boldsymbol{q}$ and $\Pi_W\psi$ denote components of the projection of $\boldsymbol{q}$ and ψ into $\boldsymbol{V}_h$ and W_h, respectively. The value of the projection on each simplex K is determined by requiring that the components satisfy the equations

$$(\boldsymbol{\Pi}_V\boldsymbol{q},\boldsymbol{v})_K = (\boldsymbol{q},\boldsymbol{v})_K \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \tag{3.1a}$$

$$(\Pi_W\psi, w)_K = (\psi, w)_K \qquad \forall\, w \in \mathcal{P}_{k-1}(K), \tag{3.1b}$$

$$\langle \boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n} + \sigma\Pi_W\psi, \mu\rangle_F = \langle \boldsymbol{q}\cdot\boldsymbol{n} + \sigma\psi, \mu\rangle_F \qquad \forall\, \mu \in \mathcal{P}_k(F), \tag{3.1c}$$

for all faces F of the simplex K.

We will also need the standard L^2-orthogonal projection onto W_h which will be denoted by P_k, and the L^2-orthogonal projection P_M onto M_h. Since σ is a piecewise constant on $\partial\Omega_h$, we have that

$$\langle \sigma(P_M\psi - \psi), \mu\rangle = 0, \quad \forall\, \mu \in M_h.$$

We will repeatedly use this fact without explicit mention. We define norms $\|\cdot\|_c$ and $\|\cdot\|_h$ as

$$\|\boldsymbol{q}\|_c^2 := \langle c\boldsymbol{q},\boldsymbol{q}\rangle, \qquad \|\psi\|_h^2 := \langle h_K\psi,\psi\rangle. \tag{3.2}$$

We are now ready to state our main result. Its proof will be given in the following section.

Theorem 3.1. *Let $y,u,z \in H^{k+2}(\Omega)$ and $\boldsymbol{p},\boldsymbol{r} \in \boldsymbol{H}^{k+2}(\Omega)$ be the solution of the optimal control system (2.1) and (2.2). Let $(\boldsymbol{p}_h, y_h, \widehat{y}_h, \boldsymbol{r}_h, z_h, \widehat{z}_h)$ be the solution obtained by the HDG method (2.3) and u_h be defined by (2.5). Then there is a constant C independent of y, u, and z such that*

$$\sqrt{\alpha}\|\boldsymbol{p}-\boldsymbol{p}_h\|_c + \|\boldsymbol{r}-\boldsymbol{r}_h\|_c \le C\left(\eta_1+\eta_2\right),$$

$$\sqrt{\alpha}\|y-y_h\| + \|z-z_h\| \le C\left(h^*\eta_1+\eta_2\right),$$

$$\|u-u_h\| \le \|P_k u - u\| + C\left(h^*\eta_1+\eta_2\right),$$

$$\|(\boldsymbol{p}-\widehat{\boldsymbol{p}}_h)\cdot\boldsymbol{n}\|_h + \|(\boldsymbol{r}-\widehat{\boldsymbol{r}}_h)\cdot\boldsymbol{n}\|_h \le \vartheta_1 + C\left(\eta_1+\eta_2\right),$$

$$\|y-\widehat{y}_h\|_h + \|z-\widehat{z}_h\|_h \le \vartheta_2 + C(\sqrt{h^*}\eta_1+\eta_2),$$

where $h^* = h^{\min(k,1)}$ and

$$\eta_1 := \sqrt{\alpha}\|\boldsymbol{p}-\boldsymbol{\Pi}_V\boldsymbol{p}\|_c + \|\boldsymbol{r}-\boldsymbol{\Pi}_V\boldsymbol{r}\|_c,$$

$$\eta_2 := \|y-\Pi_W y\| + \|z-\Pi_W z\|,$$

$$\vartheta_1 := \|(\boldsymbol{p}-P_M\boldsymbol{p})\cdot\boldsymbol{n}\|_h + \|(\boldsymbol{r}-P_M\boldsymbol{r})\cdot\boldsymbol{n}\|_h,$$

$$\vartheta_2 := \|y-P_M y\|_h + \|z-P_M z\|_h.$$

Remark 3.2. Based on Theorem 3.1 and Lemma 4.1, we can determine orders of convergence of the approximations to both fluxes and scalar variables if the value of parameter σ is given. For instance, the orders of convergence of the approximation to fluxes and scalar variables are $k+1$ when σ equals a constant that is independent of h. These rates of convergence are in agreement with the ones established in [7] for HDG methods for diffusion equations.

4. Proof of Theorem 3.1

In this section, we present a detailed proof of Theorem 3.1. We will proceed in several steps. We begin in Sec. 4.1 with stating a proposition that provides the approximation properties of the projection (3.1). In Sec. 4.2 we show that the estimate of the numerical flux depends on the interpolation error of the flux variable and the error of the scalar variable. In Sec. 4.3 we present the estimate of the scalar variable. Consequently, estimates of Theorem 3.1 follow from the triangle inequality and Lemma 4.6 which provides estimates of the projections of the errors.

4.1. **The projection error estimates.** The following lemma was established in Theorem 2.1 of [7] and provides the approximation properties of the projection operator (3.1).

Lemma 4.1. *Suppose that $k \geq 0$, $\sigma|_{\partial K}$ is nonnegative and $\sigma_K^{\max} := \max \sigma|_{\partial K} > 0$. For given functions $\boldsymbol{q} \in \boldsymbol{H}^{\ell_{\boldsymbol{q}}+1}(K)$ and $\psi \in H^{\ell_\psi+1}(K)$, we define $\boldsymbol{\Pi}_V \boldsymbol{q}$ and $\Pi_W \psi$ by the system (3.1), which is uniquely solvable for $\boldsymbol{\Pi}_V \boldsymbol{q}$ and $\Pi_W \psi$. Furthermore, there is a constant C independent of K and σ such that*

$$\|\boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}\|_K \leq C h_K^{\ell_{\boldsymbol{q}}+1} |\boldsymbol{q}|_{H^{\ell_{\boldsymbol{q}}+1}(K)} + C h_K^{\ell_\psi+1} \sigma_K^* |\psi|_{H^{\ell_\psi+1}(K)},$$
$$\|\Pi_W \psi - \psi\|_K \leq C h_K^{\ell_\psi+1} |\psi|_{H^{\ell_\psi+1}(K)} + C \frac{h_K^{\ell_{\boldsymbol{q}}+1}}{\sigma_K^{\max}} |\nabla \cdot \boldsymbol{q}|_{H^{\ell_{\boldsymbol{q}}}(K)},$$

for ℓ_ψ, $\ell_{\boldsymbol{q}}$ in $[0,k]$. Here $\sigma_K^* := \max \sigma_{\partial K \setminus F^*}$, where F^* is a face of K at which $\sigma|_{\partial K}$ is maximum.

4.2. **Flux error estimates.** Let us introduce the following notation to denote various errors of approximation. Set

$$e^{\boldsymbol{p}} = \boldsymbol{p} - \boldsymbol{p}_h, \qquad e^y = y - y_h, \qquad e^{\widehat{y}} = y - \widehat{y}_h, \qquad e^u = u - u_h,$$
$$\boldsymbol{\varepsilon}_h^{\boldsymbol{p}} = \boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}_h, \quad \varepsilon_h^y = \Pi_W y - y_h, \quad \varepsilon_h^{\widehat{y}} = P_M y - \widehat{y}_h, \quad \varepsilon_h^u = P_k u - u_h.$$

We denote $e^{\boldsymbol{r}}$, e^z, $e^{\widehat{z}}$, $\boldsymbol{\varepsilon}_h^{\boldsymbol{r}}$, ε_h^z, and $\varepsilon_h^{\widehat{z}}$ in a similar way.

From (2.2) and (2.5), it is straightforward to verify that

$$(\alpha e^u - e^z, w) = 0, \quad \forall\, w \in W_h. \tag{4.1}$$

The following lemma is in the spirit of Lemma 3.2 of [7].

Lemma 4.2. *The following two identities hold*

$$(c\boldsymbol{\varepsilon}_h^{\boldsymbol{p}}, \boldsymbol{\varepsilon}_h^{\boldsymbol{p}}) + \langle \sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), (\varepsilon_h^y - \varepsilon_h^{\widehat{y}})\rangle = (c(\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), \boldsymbol{\varepsilon}_h^{\boldsymbol{p}}) + (\beta e^z, \varepsilon_h^y), \tag{4.2a}$$
$$(c\boldsymbol{\varepsilon}_h^{\boldsymbol{r}}, \boldsymbol{\varepsilon}_h^{\boldsymbol{r}}) + \langle \sigma(\varepsilon_h^z - \varepsilon_h^{\widehat{z}}), (\varepsilon_h^z - \varepsilon_h^{\widehat{z}})\rangle = (c(\boldsymbol{\Pi}_V \boldsymbol{r} - \boldsymbol{r}), \boldsymbol{\varepsilon}_h^{\boldsymbol{r}}) - (e^y, \varepsilon_h^z). \tag{4.2b}$$

Proof. It is straightforward to verify that the projections of the errors satisfy

$$(c\boldsymbol{\varepsilon}_h^{\boldsymbol{p}}, \boldsymbol{v}) - (\varepsilon_h^y, \nabla \cdot \boldsymbol{v}) + \langle \varepsilon_h^{\widehat{y}}, \boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h \setminus \partial\Omega} = (c(\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), \boldsymbol{v}), \tag{4.3a}$$
$$-(\boldsymbol{\varepsilon}_h^{\boldsymbol{p}}, \nabla w) + \langle \widehat{\boldsymbol{\varepsilon}}_h^1 \cdot \boldsymbol{n}, w\rangle = (\beta e^z, w), \tag{4.3b}$$
$$\langle \widehat{\boldsymbol{\varepsilon}}_h^1 \cdot \boldsymbol{n}, \mu\rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.3c}$$

and

$$(c\boldsymbol{\varepsilon}_h^r, \boldsymbol{v}) - (\varepsilon_h^z, \nabla \cdot \boldsymbol{v}) + \langle \varepsilon_h^{\widehat{z}}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial\Omega_h \setminus \partial\Omega} = (c(\boldsymbol{\Pi}_V \boldsymbol{r} - \boldsymbol{r}), \boldsymbol{v}), \tag{4.4a}$$

$$- (\boldsymbol{\varepsilon}_h^p, \nabla w) + \langle \widehat{\boldsymbol{\varepsilon}}_h^2 \cdot \boldsymbol{n}, w \rangle = -(e^y, w), \tag{4.4b}$$

$$\langle \widehat{\boldsymbol{\varepsilon}}_h^2 \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega_h \setminus \partial\Omega} = 0, \tag{4.4c}$$

for all $(\boldsymbol{v}, w, \mu) \in \boldsymbol{V}_h \times W_h \times M_h$, where

$$\begin{aligned} \widehat{\boldsymbol{\varepsilon}}_h^1 \cdot \boldsymbol{n} &= \boldsymbol{\varepsilon}_h^p \cdot \boldsymbol{n} + \sigma(\varepsilon_h^y - \varepsilon^{\widehat{y}_h}) = P_M(\boldsymbol{p} \cdot \boldsymbol{n}) - \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n} \qquad \text{on } \partial\Omega_h \setminus \partial\Omega, \\ \widehat{\boldsymbol{\varepsilon}}_h^2 \cdot \boldsymbol{n} &= \boldsymbol{\varepsilon}_h^r \cdot \boldsymbol{n} + \sigma(\varepsilon_h^z - \varepsilon^{\widehat{z}_h}) = P_M(\boldsymbol{r} \cdot \boldsymbol{n}) - \widehat{\boldsymbol{r}}_h \cdot \boldsymbol{n} \qquad \text{on } \partial\Omega_h \setminus \partial\Omega. \end{aligned} \tag{4.5}$$

Taking $\boldsymbol{v} = \boldsymbol{\varepsilon}_h^p$, $w = \varepsilon_h^y$, and $\mu = -\varepsilon_h^{\widehat{y}}$ in (4.3) and adding all equations together, we obtain (4.2a) after an integration by parts. The identity (4.2b) can be proved in a similar way. □

Lemma 4.3. *We have*

$$\begin{aligned} \alpha\|\boldsymbol{\varepsilon}_h^p\|_c^2 + \|\boldsymbol{\varepsilon}_h^r\|_c^2 + \Xi &\le C\left(\sqrt{\alpha}\|\boldsymbol{p} - \boldsymbol{\Pi}_V \boldsymbol{p}\| + \|\boldsymbol{r} - \boldsymbol{\Pi}_V \boldsymbol{r}\|\right)\left(\sqrt{\alpha}\|\boldsymbol{\varepsilon}_h^p\|_c + \|\boldsymbol{\varepsilon}_h^r\|_c\right) \\ &\quad + C\left(\|y - \Pi_W y\| + \|z - \Pi_W z\|\right)\left(\|\varepsilon_h^y\| + \|\varepsilon_h^z\|\right) \end{aligned} \tag{4.6}$$

and

$$\|P_M(\boldsymbol{p} \cdot \boldsymbol{n}) - \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}\|_h + \|P_M(\boldsymbol{r} \cdot \boldsymbol{n}) - \widehat{\boldsymbol{r}}_h \cdot \boldsymbol{n}\|_h \le C_{1,\sigma}\left(\alpha\|\boldsymbol{\varepsilon}_h^p\|_c^2 + \|\boldsymbol{\varepsilon}_h^r\|_c^2 + \Xi\right) \tag{4.7}$$

where

$$\begin{aligned} \Xi &:= \alpha\langle \sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), (\varepsilon_h^y - \varepsilon_h^{\widehat{y}})\rangle + \langle \sigma(\varepsilon_h^z - \varepsilon_h^{\widehat{z}}), (\varepsilon_h^z - \varepsilon_h^{\widehat{z}})\rangle, \\ C_{1,\sigma} &:= \frac{C}{\alpha}\max\{1, h_K \sigma_K^{\max} : K \in \Omega_h\}, \end{aligned}$$

for some constant C.

Proof. Notice that

$$\begin{aligned} (\beta e^z, \varepsilon_h^y) &= (\beta(z - \Pi_W z), \varepsilon_h^y) + (\beta \varepsilon^z, \varepsilon_h^y), \\ -(e^y, \varepsilon_h^z) &= -(y - \Pi_W y, \varepsilon_h^z) - (\varepsilon_h^y, \varepsilon_h^z). \end{aligned}$$

Substituting these identities into (4.2a) and (4.2b) and adding them together, we get

$$\begin{aligned} \alpha\,(c\boldsymbol{\varepsilon}_h^p, \boldsymbol{\varepsilon}_h^p) + (c\boldsymbol{\varepsilon}_h^r, \boldsymbol{\varepsilon}_h^r) + \Xi &= \alpha\,(c(\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), \boldsymbol{\varepsilon}_h^p) + (c(\boldsymbol{\Pi}_V \boldsymbol{r} - \boldsymbol{r}), \boldsymbol{\varepsilon}_h^r) \\ &\quad - (y - \Pi_W y, \varepsilon_h^y) + (z - \Pi_W z, \varepsilon_h^y). \end{aligned} \tag{4.8}$$

Applying Cauchy-Schwarz inequality to (4.8) yields (4.6).

To prove (4.7), we apply the trace inequality to equation (4.5), which leads to

$$\begin{aligned} \|P_M(\boldsymbol{p} \cdot \boldsymbol{n}) - \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}\|_h^2 &\le \|\boldsymbol{\varepsilon}_h^p \cdot \boldsymbol{n}\|_h^2 + \|\sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}})\|_h^2 \\ &\le C\|\boldsymbol{\varepsilon}_h^p\|_c^2 + \max_{K \in \Omega_h}(h_K \sigma_K^{\max})\langle \sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), (\varepsilon_h^y - \varepsilon_h^{\widehat{y}})\rangle \\ &\le C_{1,\sigma}\left(\alpha\|\boldsymbol{\varepsilon}_h^p\|_c^2 + \alpha\langle \sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), (\varepsilon_h^y - \varepsilon_h^{\widehat{y}})\rangle\right). \end{aligned}$$

This completes the proof. □

4.3. **Estimates of scalar variables.** Consider the dual system:

$$c\mathbf{\Phi} + \nabla\Psi = 0, \qquad \text{in } \Omega, \tag{4.9a}$$

$$\nabla\cdot\mathbf{\Phi} + \psi = \Theta_1, \qquad \text{in } \Omega, \tag{4.9b}$$

$$\Psi = 0, \qquad \text{on } \partial\Omega, \tag{4.9c}$$

$$c\mathbf{\Phi}^r + \nabla\Psi^z = 0, \qquad \text{in } \Omega, \tag{4.9d}$$

$$\nabla\cdot\mathbf{\Phi}^r - \Psi = \Theta_2, \qquad \text{in } \Omega, \tag{4.9e}$$

$$\Psi^z = 0, \qquad \text{on } \partial\Omega, \tag{4.9f}$$

$$\alpha\psi - \Psi^z = 0, \qquad \text{in } \Omega. \tag{4.9g}$$

This system is equivalent to the optimal control problem:

$$\min_{\psi,\Psi} \frac{1}{2}\|\Psi + \Theta_2\|^2 + \frac{\alpha}{2}\|\psi\|^2, \tag{4.10}$$

subject to

$$\begin{aligned} -\nabla\cdot(a\nabla\Psi) &= \Theta_1 - \psi \qquad \text{in } \Omega,\\ \Psi &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{4.11}$$

Let Ψ_0 be the solution of state equation (4.11) with $\psi = 0$ and let ψ, Ψ, Ψ^z, be the optimal control, the state, and the adjoint, respectively. Then, from (4.10), we get

$$\frac{1}{2}\|\Psi + \Theta_2\|^2 + \frac{\alpha}{2}\|\psi\|^2 \le \frac{1}{2}\|\Psi_0 + \Theta_2\|^2, \tag{4.12}$$

which implies

$$\|\psi\| \le C\left(\|\Psi_0\| + \|\Theta_2\|\right). \tag{4.13}$$

We assume that the boundary value problem (4.11) admits the regularity estimate

$$\|\Psi\|_{H^2} + \|\mathbf{\Phi}\|_{H^1} \le C\|\Theta_1 - \psi\| \le C(\|\Theta_1\| + \|\psi\|). \tag{4.14}$$

This is well known to hold in several cases, for instance, if $c = 1$ and Ω is a convex polygon [13]. Consequently, when $\psi = 0$ one has

$$\|\Psi_0\|_{H^2} \le C\|\Theta_1\|. \tag{4.15}$$

Substituting (4.13) and (4.15) into (4.14) gives rise to the inequality

$$\|\Psi\|_{H^2} + \|\mathbf{\Phi}\|_{H^1} \le C\left(\|\Theta_1\| + \|\Theta_2\|\right). \tag{4.16}$$

Similarly, we can prove the regularity estimate

$$\|\Psi^z\|_{H^2} + \|\mathbf{\Phi}^r\|_{H^1} \le C\left(\|\Theta_1\| + \|\Theta_2\|\right). \tag{4.17}$$

Recall that we have been tacitly assuming that (p, y) is in the domain of Π_h. By (4.16) and (4.17), $(\mathbf{\Phi}, \Psi)$ and $(\mathbf{\Phi}^r, \Psi^z)$ are also regular enough to apply Π_h.

Next, we are going to estimate ε_h^y and ε_h^z. We start by listing two Lemmas that will be used in the proof of Lemma 4.6. The first lemma presents a weak commutativity property of operator Π_h.

Lemma 4.4. *(Proposition 2.1, [7]) For any $\omega \in W_h$ and any $(\mathbf{\Phi}, \Psi)$ in the domain of Π_h, we have*

$$(\omega, \nabla\cdot\mathbf{\Phi})_K = (\omega, \nabla\cdot\mathbf{\Pi}_V\mathbf{\Phi})_K + \langle\omega, \sigma(\Pi_W\Psi - \Psi)\rangle_{\partial K},$$

for any $K \in \mathcal{T}_h$.

Lemma 4.5. *Suppose that r, s, t are three nonnegative real numbers. If $r^2 \le rs + t^2$, then $r \le s + t$.*

Proof. We will prove that if $r > s+t$ then $r^2 > rs + t^2$. If $r > s+t$ then $r^2 > (s+t)r = rs + rt$ since r is nonnegative. Also $r > s+t$ implies that $r > t$ since s is nonnegative. Thus, $r^2 > rs + t \cdot t = rs + t^2$. □

Lemma 4.6. *Under the assumption of Theorem 3.1, there exist a constant C that is dependent of a and α, such that*

$$\sqrt{\alpha}\|\boldsymbol{\varepsilon}_h^p\|_c + \|\boldsymbol{\varepsilon}_h^r\|_c + \Xi^{\frac{1}{2}} \le C\left(\eta_1 + \eta_2\right), \tag{4.18a}$$

$$\sqrt{\alpha}\|\varepsilon_h^y\| + \|\varepsilon_h^z\| \le C\left(h^*\eta_1 + \eta_2\right), \tag{4.18b}$$

$$\|u - u_h\| \le \|P_k u - u\| + C\left(h^*\eta_1 + \eta_2\right), \tag{4.18c}$$

$$\|\varepsilon_h^{\widehat{y}}\|_h + \|\varepsilon_h^{\widehat{z}}\|_h \le C(\sqrt{h^*}\eta_1 + \eta_2), \tag{4.18d}$$

where $h^* = h^{\min(k,1)}$ and

$$\eta_1 = \sqrt{\alpha}\|\boldsymbol{p} - \boldsymbol{\Pi}_V \boldsymbol{p}\|_c + \|\boldsymbol{r} - \boldsymbol{\Pi}_V \boldsymbol{r}\|_c,$$
$$\eta_2 = \|y - \Pi_W y\| + \|z - \Pi_W z\|.$$

Proof. We will use the technique of Lemma 4.1 of [7] in the estimate of $(\varepsilon_h^y, \varepsilon_h^y)$ and $(\varepsilon_h^z, \varepsilon_h^z)$. Setting $\Theta_1 = \varepsilon_h^y$ in (4.9b), we obatin

$$\begin{aligned}
(\varepsilon_h^y, \varepsilon_h^y) &= (\varepsilon_h^y, \nabla\cdot\boldsymbol{\Phi}) + (\varepsilon_h^y, \psi)\\
&= (\varepsilon_h^y, \nabla\cdot\boldsymbol{\Phi}) + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 4.9g)}\\
&= (\varepsilon_h^y, \nabla\cdot\boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\varepsilon_h^y, \sigma(\Pi_W\Psi - \Psi)\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z) && \text{(By Lemma 4.4)}\\
&= (c\boldsymbol{\varepsilon}_h^p, \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\varepsilon_h^{\widehat{y}}, \boldsymbol{\Pi}_V\boldsymbol{\Phi}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\backslash\partial\Omega} - (c(\boldsymbol{\Pi}_V\boldsymbol{p} - \boldsymbol{p}), \boldsymbol{\Pi}_V\boldsymbol{\Phi})\\
&\quad + \langle\varepsilon_h^y, \sigma(\Pi_W\Psi - \Psi)\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 4.3a)}\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\varepsilon_h^{\widehat{y}}, (\boldsymbol{\Pi}_V\boldsymbol{\Phi} - \boldsymbol{\Phi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\backslash\partial\Omega}\\
&\quad + \langle\varepsilon_h^y, \sigma(\Pi_W\Psi - \Psi)\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z)\,,
\end{aligned}$$

by the continuity of $\boldsymbol{\Phi}\cdot\boldsymbol{n}$. Then

$$\begin{aligned}
(\varepsilon_h^y, \varepsilon_h^y) &= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \sigma(\Pi_W\Psi - \Psi)\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 3.1c)}\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), \Pi_W\Psi\rangle_{\partial\Omega_h\backslash\partial\Omega}\\
&\quad + \langle\sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), P_M\Psi\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z)\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + \langle\sigma(\varepsilon_h^y - \varepsilon_h^{\widehat{y}}), \Pi_W\Psi\rangle_{\partial\Omega_h\backslash\partial\Omega}\\
&\quad + \langle\boldsymbol{\varepsilon}_h^p\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 3.1c)}\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) - (\nabla\cdot\boldsymbol{\varepsilon}_h^p, \Pi_W\Psi) + (\beta e^z, \Pi_W\Psi)\\
&\quad + \langle\boldsymbol{\varepsilon}_h^p\cdot\boldsymbol{n}, \Psi\rangle_{\partial\Omega_h\backslash\partial\Omega} + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 4.3b)}\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) - (\nabla\cdot\boldsymbol{\varepsilon}_h^p, \Psi) + \langle\boldsymbol{\varepsilon}_h^p\cdot\boldsymbol{n}, \Psi\rangle_{\partial\Omega_h\backslash\partial\Omega}\\
&\quad + (\beta e^z, \Pi_W\Psi) + (\varepsilon_h^y, \beta\Psi^z) && \text{(By 3.1b)}\\
&= (c(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{\Pi}_V\boldsymbol{\Phi}) + (\boldsymbol{\varepsilon}_h^p, \nabla\Psi) + (\beta e^z, \Pi_W\Psi) + (\varepsilon_h^y, \beta\Psi^z)\,.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}(\varepsilon_h^y, \varepsilon_h^y) &= (c(\boldsymbol{p}-\boldsymbol{p}_h), \boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}) + (\boldsymbol{p}-\boldsymbol{p}_h, c\boldsymbol{\Phi}) + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \nabla\Psi) \\ &\quad + (\boldsymbol{p}-\boldsymbol{p}_h, \nabla\Psi) + (\beta e^z, \Pi_W\Psi) + (\varepsilon_h^y, \beta\Psi^z) \\ &= (c(\boldsymbol{p}-\boldsymbol{p}_h), \boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}) + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \nabla\Psi) \\ &\quad + (\beta e^z, \Pi_W\Psi) + (\varepsilon_h^y, \beta\Psi^z) . \qquad \text{(By 4.9a)}\end{aligned}$$

Applying (3.1a) again brings the above expression into the form

$$(\varepsilon_h^y, \varepsilon_h^y) = J_1 + (\beta e^z, \Pi_W\Psi) + (\beta\varepsilon_h^y, \Pi_W\Psi^z) + (\beta\varepsilon_h^y, \Psi^z - \Pi_W\Psi^z) , \tag{4.19}$$

where

$$J_1 = (c(\boldsymbol{p} - \boldsymbol{\Pi}_V\boldsymbol{p}), \boldsymbol{\Pi}_V\boldsymbol{\Phi} - \boldsymbol{\Phi}) + (c\boldsymbol{\varepsilon}_h^p, \boldsymbol{\Pi}_V\boldsymbol{\Phi} - \boldsymbol{\Phi}) + (\boldsymbol{p} - \boldsymbol{\Pi}_V\boldsymbol{p}, \nabla\Psi - \nabla\Psi_h) ,$$

for any $\Psi_h, \Psi_h^z \in W_h$.

Similarly, one has

$$\begin{aligned}(\varepsilon_h^z, \varepsilon_h^z) &= (\varepsilon_h^z, \nabla\cdot\boldsymbol{\Phi}^r) - (\varepsilon_h^z, \Psi) \\ &= J_2 - (e^y, \Pi_W\Psi^z) - (\varepsilon_h^z, \Pi_W\Psi) - (\varepsilon_h^z, \Psi - \Pi_W\Psi) \\ &= J_2 - (\varepsilon_h^y, \Pi_W\Psi^z) - (y - \Pi_W y, \Pi_W\Psi^z) \\ &\quad - (e^z, \Pi_W\Psi) - ((z - \Pi_W z), \Pi_W\Psi) - (\varepsilon_h^z, \Psi - \Pi_W\Psi) ,\end{aligned} \tag{4.20}$$

where

$$J_2 = (c(\boldsymbol{r} - \boldsymbol{\Pi}_V\boldsymbol{r}), \boldsymbol{\Pi}_V\boldsymbol{\Phi}^r - \boldsymbol{\Phi}^r) + (c\boldsymbol{\varepsilon}_h^r, \boldsymbol{\Pi}_V\boldsymbol{\Phi}^r - \boldsymbol{\Phi}^r) + (\boldsymbol{r} - \boldsymbol{\Pi}_V\boldsymbol{r}, \nabla\Psi^z - \nabla\Psi_h^z) ,$$

for any $\Psi_h, \Psi_h^z \in W_h$.

Multiplying (4.20) by α, adding it to (4.19), and setting $w = \Pi_W\Psi$ in (4.1), we obtain

$$\begin{aligned}\alpha\,(\varepsilon_h^y, \varepsilon_h^y) + (\varepsilon_h^z, \varepsilon_h^z) &= \alpha J_1 + J_2 + (\varepsilon_h^y, \Psi^z - \Pi_W\Psi^z) - (y - \Pi_W y, \Pi_W\Psi^z) \\ &\quad - ((z - \Pi_W z), \Pi_W\Psi) - (\varepsilon_h^z, \Psi - \Pi_W\Psi) .\end{aligned}$$

It follows from Cauchy-Schwarz inequality and Lemma 4.1 that

$$\alpha\|\varepsilon_h^y\|^2 + \|\varepsilon_h^z\|^2 \le C\zeta\left(h^*\left(\eta_1 + \theta_1\right) + \eta_2 + h^*\theta_2\right),$$

where

$$\begin{aligned}\zeta &:= \sqrt{\alpha}\|\boldsymbol{\Phi}\|_{H^1} + \sqrt{\alpha}\|\Psi\|_{H^1} + \|\boldsymbol{\Phi}^r\|_{H^1} + \|\Psi^z\|_{H^1}, \\ \theta_1 &:= \sqrt{\alpha}\|\boldsymbol{\varepsilon}_h^p\|_c + \|\boldsymbol{\varepsilon}_h^r\|_c, \\ \theta_2 &:= \sqrt{\alpha}\|\varepsilon_h^y\| + \|\varepsilon_h^z\|.\end{aligned}$$

If h is sufficiently small, it follows from (4.16) and (4.17) with $\Theta_1 = \varepsilon_h^y$ and $\Theta_2 = \varepsilon_h^z$ that

$$\sqrt{\alpha}\|\varepsilon_h^y\| + \|\varepsilon_h^z\| \le C_1 h^*\left(\eta_1 + \theta_1\right) + C_2\eta_2, \tag{4.21}$$

for some constants C_1 and C_2. Substituting (4.21) into (4.6) we obtain

$$\begin{aligned}\theta_1^2 + \Xi = \alpha\|\boldsymbol{\varepsilon}_h^p\|_c^2 + \|\boldsymbol{\varepsilon}_h^r\|_c^2 + \Xi &\le C\left[\eta_1\theta_1 + h^*\eta_2\left(\eta_1 + \theta_1\right) + \eta_2^2\right] \\ &\le C_1\eta_1\theta_1 + C_2\left(h^*\eta_2\eta_1 + \eta_2^2\right) + \frac{1}{2}\theta_1^2, \\ &\le C_1\eta_1\theta_1 + C_2\left(h^*\eta_1^2 + \eta_2^2\right) + \frac{1}{2}\theta_1^2.\end{aligned}$$

Moving the term $\frac{1}{2}\theta_1^2$ of the above inequality from the right side to the left side and applying Lemma 4.5, we obtain

$$\theta_1 + \Xi^{\frac{1}{2}} \leq C\left(\eta_1 + \eta_2\right).$$

This completes the proof of (4.18a). The estimate (4.18b) is established by substituting the above inequality into (4.21).

To estimate e^u, we note that

$$\begin{aligned}\|\varepsilon_h^u\|^2 = (\varepsilon_h^u, \varepsilon_h^u) &= (P_k u - u, \varepsilon_h^u) + (e^u, \varepsilon_h^u)\\ &= (P_k u - u, \varepsilon_h^u) + (\beta e^z, \varepsilon_h^u)\\ &\leq \left(\|P_k u - u\| + \beta\|e^z\|\right)\|\varepsilon_h^u\|,\end{aligned}$$

by the equation (4.1). Then, using (4.18b) we get that

$$\begin{aligned}\|e_u\| &\leq \|P_k u - u\| + \|\varepsilon_h^u\|\\ &\leq \|P_k u - u\| + \beta\|z - \Pi_W z\| + C\left(h^*\eta_1 + \eta_2\right)\\ &\leq \|P_k u - u\| + C\left(h^*\eta_1 + \eta_2\right).\end{aligned}$$

The estimate of $\varepsilon_h^{\widehat{y}}$ follows from the same local argument used in [3] to obtain a similar estimate for the BDM method. Indeed, when $k \geq 1$, we can select a function $\boldsymbol{r} \in \boldsymbol{\mathcal{P}}_k(K)$ such that $\boldsymbol{r}\cdot\boldsymbol{n} = \varepsilon_h^{\widehat{y}}$ on ∂K and $\|\boldsymbol{r}\|_K \leq C h_K^{\frac{1}{2}}\|\varepsilon_h^{\widehat{y}}\|_{\partial K}$. Using $h_K\boldsymbol{r}$ as the test function in (4.3), and applying an inverse inequality, we obtain

$$\begin{aligned}h_K\|\varepsilon_h^{\widehat{y}}\|_{\partial K}^2 &= h_K(c(\boldsymbol{\Pi}_V\boldsymbol{p} - \boldsymbol{p}), \boldsymbol{r}) - h_K\left(c\boldsymbol{\varepsilon}_h^p, \boldsymbol{r}\right) + h_K\left(\varepsilon_h^y, \nabla\cdot\boldsymbol{r}\right)\\ &\leq C h_K\|\boldsymbol{r}\|_K\left(\|\boldsymbol{\Pi}_V\boldsymbol{p} - \boldsymbol{p}\|_K + \|\boldsymbol{\varepsilon}_h^p\|_K\right) + C\|\boldsymbol{r}\|_K\|\varepsilon_h^y\|_K\\ &\leq C\left[h_K\left(\|\boldsymbol{\Pi}_V\boldsymbol{p} - \boldsymbol{p}\|_K + \|\boldsymbol{\varepsilon}_h^p\|_K\right)^2 + \|\varepsilon_h^y\|_K^2\right] + \frac{1}{2}h_K\|\varepsilon_h^{\widehat{y}}\|_{\partial K}^2.\end{aligned}$$

Canceling the term $\frac{1}{2}h_K\|\varepsilon_h^{\widehat{y}}\|_{\partial K}^2$ yields

$$h_K^{1/2}\|\varepsilon_h^{\widehat{y}}\|_{\partial K} \leq C\left(h_K^{1/2}\|\boldsymbol{\Pi}_V\boldsymbol{p} - \boldsymbol{p}\|_K + h_K^{1/2}\|\boldsymbol{\varepsilon}_h^p\|_K + \|\varepsilon_h^y\|_K\right).$$

Applying (4.18a) and (4.18b), we get the last inequality of the theorem. This completes the proof. □

5. Conclusion

In this paper, we derived a priori error estimates of an HDG method for an optimal control problem governed by a second order elliptic partial differential e-quation. Optimal orders of convergence can be obtained for suitably chosen values of the parameter σ. This is the first stepping stone for devising HDG methods for more general optimal control problems. The next natural step is to study HDG methods for optimal control problems governed by convection-dominated PDEs, which is the subject of ongoing work. Implementing these methods and validating our theoretical estimates are also under investigation and will be published elsewhere.

Acknowledgements: The second author was partially supported by the National Science Foundation (Grant DMS-1115280). The authors thank the anonymous referees for their critical and constructive comments which improved the presentation and the quality of this paper.

References

[1] S.R. Arridge, *Optimal tomography in medical imaging*, Inverse Problems, 15:R41-R93, 1999.

[2] I. Bouchouev and V. Isakov, *Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets*, Inverse Problems, 15:R95-R116, 1999.

[3] F. Brezzi, J. Dauglas, JR., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47, 217–235.

[4] Y. Chen and B. Cockburn, *Analysis of variable-degree HDG methods for convection-diffusion equations. Part I: general nonconforming meshes*, IMA Journal of Numerical Analysis, doi:10.1093/imanum/drr058

[5] Y. Chen, Y. Huang, W.B. Liu, and N.N. Yan, *Error estimates and superconvergence of mixed finite element methods for convex optimal control problems*, J. Sci. Comput., 42: 382–403, 2010.

[6] B. Cockburn, J. Gopalakrishnan, and R. Lazarov *Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. 47(2), 1319–1365.

[7] B. Cockburn, J. Gopalakrishnan, and F. Sayas, A projection-based error analysis of HDG methods Math. Comp. 79, 1351–1367, 2010.

[8] S.S. Collis and M. Heinkenschloss, *Numerical solution of implicitly constrained optimization problems.* Technical Report Technical Report TR08/05, Department of Computational and Applied Mathematics, Rice University, Houston, 2008. http://www.caam.rice.edu/ heinken.

[9] H. Egger and H. W. Engl, *Tikhonov regularization applied to the inverse problem of option pricing: Convergence analysis and rates*, Inverse Problems, 21:1027-1045, 2005.

[10] O. Ghattas and C. Orozco, *Massively parallel aerodynamic shape optimization*, Comput. Syst. Eng., 1:311-320, 1992.

[11] W. Gong, R. Li, N.N. Yan, W.B. Zhao, *An improved error analysis for finite element approximation of bioluminescence tomography*, J. Comput. Math. 26(3), 297-309 (2008).

[12] W. Gong, N.N. Yan, *A posteriori error estimate for boundary control problems governed by the parabolic partial differential equations.* J. Comput. Math. 27(1), 68-88, 2009.

[13] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, no. 24 in Monographs and Studies in Mathe- matics, Pitman Advanced Publishing Program, Marshfield, Massachusetts, 1985.

[14] M.V. Klibanov and T.R. Lucas, *Numerical solution of a parabolic inverse problem in optical tomography using experimental data*, SIAM J. Appl. Math., 59:6516-6534, 1999.

[15] D. Leykekhman, *Investigation of commutative properties of discontinuous Galerkin methods in PDE constrained optimal control problems*, J. Sci. Comput., 53(3): 483-511, 2012.

[16] J. L. Lions, *Optimal Control of Systems*, Springer, 1968.

[17] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer Verlag, Berlin, Heidelberg, New York, 1971.

[18] J. L. Lions and E. Magenes, *Non-homogeneous boundary values and applications*, Springer-Verlag, Berlin, 1972.

[19] W.B. Liu, N.N. Yan, *A posteriori error estimates for optimal problems governed by Stokes equations*, SIAM J. Numer. Anal. 40, 1850-1869, 2003.

[20] W.B. Liu, H. Ma, T. Tang, and N.N. Yan, *A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*, SIAM J. Numer. Anal. 42(3): 1032–1061, 2004.

[21] A. Quarteroni, *Numerical models for differential problems.* Springer-Verlag Italia, Milan, 2009.

[22] T. Rees, *Preconditioning iterative methods for PDE constrained optimization*, Ph.D. thesis, University of Oxford, 2010.

[23] A. Shenoy, M. Heinkenschloss, and E. M. Cliff, *Airfoil design by an all-at-once method*, Int. J. Comput. Fluid Dyn., 11:3-25, 1998.

[24] N.N. Yan and W.B. Liu, *Adaptive finite element methods for optimal control governed by PDEs*, Science Press, 2008.

[25] H. Yücel, M. Heinkenschloss, and B. Karasözen, *Distributed optimal control of diffusion-convection-reaction equations using discontinuous Galerkin methods*, Numerical Mathematics and Advanced Applications 2011 2013, pp 389–397.

[26] Z. Zhou and N.N. Yan, *The local discontinuous Galerkin method for optimal control problem governed by convection-diffusion equations*, Int. J. Numer. Anal. Model. 7(4), 681–699, 2010.

Department of Mathematics, University of Southern Mississippi, Hattiesburg, MS 39406, USA.
E-mail address: Huiqing.Zhu@usm.edu

Department of Mathematics, Wayne State University, Detroit, MI 48202, USA.
E-mail address: celiker@math.wayne.edu