SIAM J. NUMER. ANAL.
Vol. 57, No. 4, pp. 1919–1946

AN HDG METHOD FOR DIRICHLET BOUNDARY CONTROL OF CONVECTION DOMINATED DIFFUSION PDEs*

GANG CHEN†, JOHN R. SINGLER‡, AND YANGWEN ZHANG§

Abstract. We first propose a hybridizable discontinuous Galerkin (HDG) method to approximate the solution of a *convection dominated* Dirichlet boundary control problem without constraints. Dirichlet boundary control problems and convection dominated problems are each very challenging numerically due to solutions with low regularity and sharp layers, respectively. Although there are some numerical analysis works in the literature on *diffusion dominated* convection diffusion Dirichlet boundary control problems, we are not aware of any existing numerical analysis works for convection dominated boundary control problems. Moreover, the existing numerical analysis techniques for convection dominated PDEs are not directly applicable for the Dirichlet boundary control problem because of the low regularity solutions. In this work, we obtain an optimal a priori error estimate for the control under some conditions on the domain and the desired state. We also present some numerical experiments to illustrate the performance of the HDG method for convection dominated Dirichlet boundary control problems.

Key words. Dirichlet boundary control, convection dominated diffusion PDEs, error analysis, low regularity, hybridizable discontinuous Galerkin method, HDG method

AMS subject classifications. 49J20, 65M60

DOI. 10.1137/18M1208708

1. Introduction.

Let $\Omega \subset \mathbb{R}^2$ be a Lipschitz polygonal domain with boundary $\Gamma = \partial\Omega$. We consider the following unconstrained Dirichlet boundary control problem,

$$\min J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u\|^2_{L^2(\Gamma)}, \quad \gamma > 0, \tag{1.1}$$

subject to

$$\begin{aligned} -\varepsilon\Delta y + \nabla\cdot(\boldsymbol{\beta} y) + \sigma y &= f \quad \text{in } \Omega,\\ y &= u \quad \text{on } \Gamma, \end{aligned} \tag{1.2}$$

where $f \in L^2(\Omega), \varepsilon \ll |\boldsymbol{\beta}|$, and we make other assumptions on $\boldsymbol{\beta}$ and σ for our analysis.

Researchers have performed numerical analysis of computational methods for Dirichlet boundary control problems for over a decade. Many researchers considered

*Received by the editors August 21, 2018; accepted for publication (in revised form) May 21, 2019; published electronically August 13, 2019.
https://doi.org/10.1137/18M1208708
Funding: The work of the first author was supported by the National Natural Science Foundation of China grant 11801063 and the China Postdoctoral Science Foundation grant 2018M633339. The work of the second and third authors was partially supported by the National Science Foundation grant DMS-1217122. The second and third authors also received funding from the Institute for Mathematics and its Applications at the University of Minnesota United States for research visits, during which some of this work was completed. The work of the third author was partially supported by the National Science Foundation grant DMS-1619904.
†School of Mathematics, Sichuan University, Chengdu, China and School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China (cglwdm@scu.edu.cn).
‡Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO 65409 (singlerj@mst.edu).
§Department of Mathematics Science, University of Delaware, Newark, DE 19716 (ywzhangf@udel.edu).

the standard finite element method and obtained an error estimate for the constrained Dirichlet boundary control of order h^s for all $s < \min\{1, \pi/2\omega\}$, where ω is the largest angle of the boundary polygon (see, e.g., [9, 40, 39]). Apel et al. in [2] considered special meshes and obtained an optimal convergence rate with $s < \min\{3/2, \pi/\omega - 1/2\}$. Some mixed finite element methods have also been used for Dirichlet boundary control problems because the essential Dirichlet boundary condition becomes natural, i.e., the Dirichlet boundary data directly enters the variational setting. In [26], Gong and Yan used a standard mixed method to obtain an error estimate for all $s < \min\{1, \pi/2\omega\}$. Recently, we used a hybridizable discontinuous Galerkin (HDG) method to obtain an optimal convergence rate for all $s < \min\{3/2, \pi/\omega - 1/2\}$ without using higher order elements or a special mesh [33]. Moreover, the number of degrees of freedom are lower for HDG methods than standard mixed methods.

All of the above works focus on Dirichlet boundary control of the Poisson equation. However, Dirichlet boundary control problems play an important role in many applications governed by more complicated models, such as the Navier–Stokes equations; see, e.g., [28, 27, 29, 30, 24]. In order to work toward numerical analysis results for more difficult PDEs, one essential and necessary step is to fully understand the convection diffusion Dirichlet boundary control problem. Benner and Yücel in [5] used a local discontinuous Galerkin (LDG) method and they obtained an error estimate for the control of order $\mathcal{O}(h^s)$ for all $s < \min\{1, \pi/2\omega\}$. Also, very recently, we proposed a new HDG method to study this problem and obtained an optimal convergence rate $\mathcal{O}(h^s)$ for all $s < \min\{3/2, \pi/\omega - 1/2\}$; see [32, 25] for more details.

However, the previous works only approximated solutions of convection diffusion Dirichlet boundary control problems in the *diffusion dominated* case. They did not consider the more difficult *convection dominated* case, i.e., $\varepsilon \ll |\boldsymbol{\beta}|$. Even without the Dirichlet boundary control, solutions of convection dominated diffusion PDEs typically have layers; therefore, designing a robust numerical scheme for this problem is a major difficulty and has been considered in many works; see, e.g., [21, 35, 41, 8] and the references therein. Discontinuous Galerkin (DG) methods have proved very useful for solving convection dominated PDEs; see, e.g., [11, 15, 7, 20, 14, 36, 10] for standard DG methods and [34, 23] for HDG methods. For more information on HDG methods, see, e.g., [16, 17, 12, 13, 19, 18, 42, 43]. Moreover, there are some existing convection dominated diffusion *distributed* optimal control numerical analysis works; see, e.g., [4, 37, 31]. However, the techniques in the above works are *not* applicable for convection dominated Dirichlet boundary control problems since the solutions of (1.1)–(1.2) frequently have low regularity, i.e., $y \in H^{1+s}(\Omega)$ with $0 \leq s < 1/2$.

Formally, the optimal control $u \in L^2(\Gamma)$ and the optimal state $y \in L^2(\Omega)$ minimizing the cost functional satisfy a mixed weak formulation of the optimality system

$$-\varepsilon\Delta y + \nabla\cdot(\boldsymbol{\beta} y) + \sigma y = f \quad \text{in } \Omega, \tag{1.3a}$$

$$y = u \quad \text{on } \Gamma, \tag{1.3b}$$

$$-\varepsilon\Delta z - \nabla\cdot(\boldsymbol{\beta} z) + (\nabla\cdot\boldsymbol{\beta} + \sigma) z = y - y_d \quad \text{in } \Omega, \tag{1.3c}$$

$$z = 0 \quad \text{on } \Gamma, \tag{1.3d}$$

$$\gamma u - \varepsilon\nabla z\cdot\boldsymbol{n} = 0 \quad \text{on } \Gamma. \tag{1.3e}$$

In this work, we use polynomials of degree k to approximate the state y, dual state z, and their fluxes $\boldsymbol{q} = -\varepsilon\nabla y$ and $\boldsymbol{p} = -\varepsilon\nabla z$, respectively. Moreover, we also use polynomials of degree k to approximate the numerical trace of the state and dual state on the edges of the spatial mesh, which are the only globally coupled unknowns. The HDG method considered here is different from the HDG method we considered for convection diffusion Dirichlet boundary control problems in [32, 25]. A major difference is that the HDG method here has a lower computational cost.

In section 4, we obtain an optimal convergence rate for the optimal control in two dimensions under certain basic assumptions on the desired state y_d and the domain Ω; specifically, we prove

$$\|u - u_h\|_\Gamma \le Ch^s \tag{1.4}$$

for all $s < \min\{3/2, \pi/\omega - 1/2\}$, and the constant C only depends on the exact solution, the domain, and the polynomial degree. To prove the estimate (1.4), we cannot use the numerical analysis strategy from [5, 32, 25] because the constants in their error estimates may blow up as ε approaches zero. In order to obtain the estimate (1.4) with the constant C independent of ε, we follow a strategy from [23] and use weighted test functions in an energy argument. However, the techniques used in [23] are not directly applicable for solutions with low regularity. Moreover, unlike all the previous Dirichlet boundary control numerical analysis works, we only assume the mesh is shape regular, not quasi-uniform. We present numerical results in section 5 to illustrate the performance of the HDG method.

2. Optimality system, regularity, and HDG formulation.

We begin with some notation. For any bounded domain $\Lambda \subset \mathbb{R}^2$, let $H^m(\Lambda)$ and $H_0^m(\Lambda)$ denote the usual mth-order Sobolev spaces on Λ, and let $\|\cdot\|_{m,\Lambda}$, $|\cdot|_{m,\Lambda}$ denote the norm and seminorm on these spaces. We use $(\cdot,\cdot)_{m,\Lambda}$ to denote the inner product on $H^m(\Lambda)$ and set $(\cdot,\cdot)_\Lambda := (\cdot,\cdot)_{0,\Lambda}$. When $\Lambda = \Omega$, we denote $\|\cdot\|_m := \|\cdot\|_{m,\Omega}$ and $|\cdot|_m := |\cdot|_{m,\Omega}$. Also, when Λ is the boundary of a set in $\mathbb{R}^2$, we use $\langle\cdot,\cdot\rangle_\Lambda$ to replace $(\cdot,\cdot)_\Lambda$. Boldface fonts will be used for vector Sobolev spaces along with vector-valued functions. In addition, we introduce the following space:

$$\boldsymbol{H}(\mathrm{div},\Lambda) := \{\boldsymbol{v} \in [L^2(\Lambda)]^2 : \nabla\cdot\boldsymbol{v} \in L^2(\Lambda)\}.$$

We now present the optimality system for problem (1.1)–(1.2) and give a regularity result.

2.1. Optimality system and regularity.

Throughout the paper, we suppose Ω is a convex polygonal domain, and let $\omega \in [\pi/3, \pi)$ denote its largest interior angle. The optimal control u is determined by the optimality system for the state y and the dual state z. For the HDG method, we use a mixed formulation of the optimality system; therefore we introduce the primary flux $\boldsymbol{q} = -\varepsilon\nabla y$ and the dual flux $\boldsymbol{p} = -\varepsilon\nabla z$. The well-posedness and regularity of the mixed formulation of the optimality system is contained in the result below. The proof of Theorem 2.1 is omitted here since it is very similar with a proof of a similar result in [32].

THEOREM 2.1. *If $y_d \in H^{t^*}(\Omega)$ for some $0 \le t^* < 1$, $\sigma \in L^\infty(\Omega)\cap H^1(\Omega)$, $f = 0$ and the velocity vector field $\boldsymbol{\beta}$ satisfies*

$$\boldsymbol{\beta} \in [L^\infty(\Omega)]^2, \quad \nabla\cdot\boldsymbol{\beta} \in L^\infty(\Omega), \quad \sigma + \frac{1}{2}\nabla\cdot\boldsymbol{\beta} \ge 0, \quad \nabla\nabla\cdot\boldsymbol{\beta} \in [L^2(\Omega)]^2, \tag{2.1}$$

then problem (1.1)–(1.2) has a unique solution $u \in L^2(\Gamma)$. Moreover, for any $s > 0$ satisfying $s \le 1/2 + t^*$ and $s < \min\{3/2, \pi/\omega - 1/2\}$, we have that

$$(u, \boldsymbol{q}, \boldsymbol{p}, y, z) \in H^s(\Gamma) \times \boldsymbol{H}^{s-\frac{1}{2}}(\Omega) \times \boldsymbol{H}^{s+\frac{1}{2}}(\Omega) \times H^{s+\frac{1}{2}}(\Omega) \times (H^{s+\frac{3}{2}}(\Omega) \cap H_0^1(\Omega))$$

is the unique solution of

$$\varepsilon^{-1}(\boldsymbol{q},\boldsymbol{r})_\Omega-(y,\nabla\cdot\boldsymbol{r})_\Omega+\langle u,\boldsymbol{r}\cdot\boldsymbol{n}\rangle_\Gamma=0, \tag{2.2a}$$

$$(\nabla\cdot(\boldsymbol{q}+\boldsymbol{\beta}y),w)_\Omega+(\sigma y,w)_\Omega=0, \tag{2.2b}$$

$$\varepsilon^{-1}(\boldsymbol{p},\boldsymbol{r})_\Omega-(z,\nabla\cdot\boldsymbol{r})_\Omega=0, \tag{2.2c}$$

$$(\nabla\cdot(\boldsymbol{p}-\boldsymbol{\beta}z),w)_\Omega+((\nabla\cdot\boldsymbol{\beta}+\sigma)y,w)_\Omega=(y-y_d,w)_\Omega, \tag{2.2d}$$

$$\langle\gamma u+\boldsymbol{p}\cdot\boldsymbol{n},v\rangle_\Gamma=0 \tag{2.2e}$$

for all $(\boldsymbol{r},w,v)\in\boldsymbol{H}(\mathrm{div},\Omega)\times L^2(\Omega)\times L^2(\Gamma)$. *Furthermore, we have* $\Delta y\in L^2(\Omega)$.

It is worth mentioning that the forcing f is identically zero in Theorem 2.1; if this is not the case, then a simple change of variable as in [1, p. 3623] can be used to eliminate the forcing. In the following error analysis, we keep the forcing f nonzero.

Moreover, we will constantly use the following lemma in our analysis; the proof is very similar in [32, Lemma 2.4], and hence we omit it.

LEMMA 2.2. *Let s be a real number such that $-1/2\le s<3/2$ and $s>1/2-\pi/\omega$. For every $u\in H^s(\Gamma)$, there exists a unique* very weak *solution $y\in H^{1/2+s}(\Omega)$ of* (1.2) *and*

$$\|y\|_{H^{1/2+s}(\Omega)}\le C\|u\|_{H^s(\Gamma)}.$$

2.2. The HDG formulation. Let $\mathcal{T}_h=\bigcup\{T\}$ be a conforming simplex mesh that partitions the domain Ω. For any $T\in\mathcal{T}_h$, we let h_T be the diameter of T and denote the mesh size by $h:=\max_{T\in\mathcal{T}_h}h_T$. Denote the edges of T by E, let $\mathcal{E}_h$ be the set of all edges E, let $\mathcal{E}_h^\partial$ be the set of edges E such that $E\subset\Gamma$, and set $\mathcal{E}_h^o=\mathcal{E}_h\setminus\mathcal{E}_h^\partial$. Let h_E denote the diameter of E. The mesh dependent inner products are denoted by

$$(w,v)_{\mathcal{T}_h}=\sum_{T\in\mathcal{T}_h}(w,v)_T,\qquad \langle\zeta,\rho\rangle_{\partial\mathcal{T}_h}=\sum_{T\in\mathcal{T}_h}\langle\zeta,\rho\rangle_{\partial T}.$$

We use ∇ and $\nabla\cdot$ to denote the broken gradient and broken divergence with respect to $\mathcal{T}_h$. For an integer $k\ge0$, $\mathcal{P}_k(\Lambda)$ denotes the set of all polynomials defined on Λ with degree not greater than k. We introduce the discontinuous finite element spaces:

$$\begin{aligned}
\boldsymbol{V}_h&:=\{\boldsymbol{v}\in[L^2(\Omega)]^2:\boldsymbol{v}|_T\in[\mathcal{P}_k(T)]^2\ \forall T\in\mathcal{T}_h\},\\
W_h&:=\{w\in L^2(\Omega):w|_T\in\mathcal{P}_k(T)\ \forall T\in\mathcal{T}_h\},\\
M_h^o&:=\{\mu\in L^2(\mathcal{E}_h):\mu|_E\in\mathcal{P}_k(E)\ \forall E\in\mathcal{E}_h,\ \text{and }\mu|_\Gamma=0\},\\
M_h^\partial&:=\{\mu\in L^2(\mathcal{E}_h^\partial):\mu|_E\in\mathcal{P}_k(E)\ \forall E\in\mathcal{E}_h^\partial\}.
\end{aligned}$$

In our earlier works [32, 25], we used a $\mathcal{P}_{k+1}$ local space for the spaces W_h and M_h. In this work, we use polynomial degree k for all spaces. Since the globally coupled degrees of freedom depend on the space M_h, the computational cost of the HDG method in this paper is much lower than the HDG method in [32, 25].

The HDG method for the mixed weak form of the optimality system (2.2) is to find $(\boldsymbol{q}_h,y_h,\widehat{y}_h^o,\boldsymbol{p}_h,z_h,\widehat{z}_h^o,u_h)\in[\boldsymbol{V}_h\times W_h\times M_h^o]^2\times M_h^\partial$ such that

$$\varepsilon^{-1}(\boldsymbol{q}_h,\boldsymbol{r}_h)_{\mathcal{T}_h}-(y_h,\nabla\cdot\boldsymbol{r}_h)_{\mathcal{T}_h}+\langle\widehat{y}_h^o,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}=-\langle u_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial} \tag{2.3a}$$

for all $\boldsymbol{r}_h\in\boldsymbol{V}_h$,

$$\begin{aligned}
&-(w_h,\nabla\cdot\boldsymbol{q}_h)_{\mathcal{T}_h}+\langle\widehat{w}_h^o,\boldsymbol{q}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}-\langle\tau_1(y_h-\widehat{y}_h^o),w_h-\widehat{w}_h^o\rangle_{\partial\mathcal{T}_h}\\
&+(y_h,\boldsymbol{\beta}\cdot\nabla w_h)_{\mathcal{T}_h}-\langle\widehat{y}_h^o,\boldsymbol{\beta}\cdot\boldsymbol{n}w_h\rangle_{\partial\mathcal{T}_h}-(\sigma y_h,w_h)_{\mathcal{T}_h}\\
&=-(f,w_h)_{\mathcal{T}_h}-\langle(\tau_1-\boldsymbol{\beta}\cdot\boldsymbol{n})u_h,v_h\rangle_{\mathcal{E}_h^\partial}
\end{aligned} \tag{2.3b}$$

for all $(w_h, \widehat{w}_h^o) \in W_h \times M_h^o$,

$$\varepsilon^{-1}(\boldsymbol{p}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (z_h, \nabla \cdot \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0 \tag{2.3c}$$

for all $r_h \in \boldsymbol{V}_h$,

$$\begin{aligned}
&- (w_h, \nabla \cdot \boldsymbol{p}_h)_{\mathcal{T}_h} + \langle \widehat{w}_h^o, \boldsymbol{p}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \tau_2 (z_h - \widehat{z}_h^o)_{\mathcal{T}_h}, w_h - \widehat{w}_h^o \rangle_{\partial \mathcal{T}_h} \\
&- (z_h, \boldsymbol{\beta} \cdot \nabla w_h)_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{\beta} \cdot \boldsymbol{n} w_h \rangle_{\partial \mathcal{T}_h} - ((\sigma + \nabla \cdot \boldsymbol{\beta}) z_h, w_h)_{\mathcal{T}_h} \\
&= -(y_h - y_d, w_h)_{\mathcal{T}_h}
\end{aligned} \tag{2.3d}$$

for all $(w_h, \widehat{w}_h^o) \in W_h \times M_h^o$, and

$$\langle \gamma u_h + \boldsymbol{p}_h \cdot \boldsymbol{n} + \tau_2 (z_h - \widehat{z}_h^o), \widehat{w}_h^\partial \rangle_{\varepsilon_h^\partial} = 0 \tag{2.3e}$$

for all $\widehat{w}_h^\partial \in M_h^\partial$. Here, the positive stabilization functions τ_1 and τ_2 are chosen as

$$\tau_1|_{\partial T} = \|\boldsymbol{\beta} \cdot \boldsymbol{n}\|_{0,\infty,\partial T} + \frac{1}{2} \boldsymbol{\beta} \cdot \boldsymbol{n} + \varepsilon h_T^{-1}, \tag{2.3f}$$

$$\tau_2|_{\partial T} = \|\boldsymbol{\beta} \cdot \boldsymbol{n}\|_{0,\infty,\partial T} - \frac{1}{2} \boldsymbol{\beta} \cdot \boldsymbol{n} + \varepsilon h_T^{-1}. \tag{2.3g}$$

To simplify the presentation later, we define

$$\tau = \frac{\tau_1 + \tau_2}{2} = \|\boldsymbol{\beta} \cdot \boldsymbol{n}\|_{0,\infty,\partial T} + \varepsilon h_T^{-1}. \tag{2.3h}$$

2.3. A compact formulation. To simplify the notation, for $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{p}_h, z_h, \widehat{z}_h^o, \boldsymbol{r}_h, w_h, \widehat{w}_h^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^3$, we denote

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o) \\
&\quad = \varepsilon^{-1}(\boldsymbol{q}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (y_h, \nabla \cdot \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&\qquad - (w_h, \nabla \cdot \boldsymbol{q}_h)_{\mathcal{T}_h} + \langle \widehat{w}_h^o, \boldsymbol{q}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \tau_1 (y_h - \widehat{y}_h^o), w_h - \widehat{w}_h^o \rangle_{\partial \mathcal{T}_h} \\
&\qquad + (y_h, \boldsymbol{\beta} \cdot \nabla w_h)_{\mathcal{T}_h} - \langle \widehat{y}_h^o, \boldsymbol{\beta} \cdot \boldsymbol{n} w_h \rangle_{\partial \mathcal{T}_h} - (\sigma y_h, w_h)_{\mathcal{T}_h},
\end{aligned} \tag{2.4}$$

$$\begin{aligned}
&\mathcal{B}_2(\boldsymbol{p}_h, z_h, \widehat{z}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o) \\
&\quad = \varepsilon^{-1}(\boldsymbol{p}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (z_h, \nabla \cdot \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&\qquad - (w_h, \nabla \cdot \boldsymbol{p}_h)_{\mathcal{T}_h} + \langle \widehat{w}_h^o, \boldsymbol{p}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \tau_2 (z_h - \widehat{z}_h^o), w_h - \widehat{w}_h^o \rangle_{\partial \mathcal{T}_h} \\
&\qquad - (z_h, \boldsymbol{\beta} \cdot \nabla w_h)_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{\beta} \cdot \boldsymbol{n} w_h \rangle_{\partial \mathcal{T}_h} - ((\sigma + \nabla \cdot \boldsymbol{\beta}) z_h, w_h)_{\mathcal{T}_h}.
\end{aligned} \tag{2.5}$$

Then we can rewrite (2.3) as follows: find $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{p}_h, z_h, \widehat{z}_h^o, u_h) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2 \times M_h^\partial$ such that

$$\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_1, w_1, \widehat{w}_1^o) = -(f, w_1)_{\mathcal{T}_h} - \langle u_h, \tau_2 w_1 + \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\mathcal{E}_h^\partial}, \tag{2.6a}$$

$$\mathcal{B}_2(\boldsymbol{p}_h, z_h, \widehat{z}_h^o; \boldsymbol{r}_2, w_2, \widehat{w}_2^o) = -(y_h - y_d, w_2)_{\mathcal{T}_h}, \tag{2.6b}$$

$$\langle \boldsymbol{p}_h \cdot \boldsymbol{n} + \tau_2 (z_h - \widehat{z}_h^o), \widehat{w}_h^\partial \rangle_{\varepsilon_h^\partial} = \langle \gamma u_h, \widehat{w}_h^\partial \rangle_{\varepsilon_h^\partial} \tag{2.6c}$$

for all $\left(\boldsymbol{r}_1, w_1, \widehat{w}_1^o, \boldsymbol{r}_2, w_2, \widehat{w}_2^o, \widehat{w}_h^\partial\right) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2 \times M_h^\partial$.

The following basic result, which is similar to results in [32, 25], is crucial to the proof of the well-posedness of the discrete optimality system (2.3a)–(2.3e) and is also a very important part of the final stage of the numerical analysis (see the proof of Lemma 4.9).

LEMMA 2.3. *For all* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{r}_h, w_h, \widehat{w}_h^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$, *we have*

$$\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o) = \mathcal{B}_2(\boldsymbol{r}_h, w_h, \widehat{w}_h^o; \boldsymbol{q}_h, y_h, \widehat{y}_h^o). \tag{2.7}$$

Proof. Using the definitions in (2.4)–(2.5) and integration by parts gives

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o) - \mathcal{B}_2(\boldsymbol{r}_h, w_h, \widehat{w}_h^o; \boldsymbol{q}_h, y_h, \widehat{y}_h^o)\\
&\quad = -\langle \tau_1(y_h - \widehat{y}_h^o), w_h - \widehat{w}_h^o \rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle \tau_2(y_h - \widehat{y}_h^o), w_h - \widehat{w}_h^o \rangle_{\partial\mathcal{T}_h}\\
&\qquad + (y_h, \boldsymbol{\beta}\cdot\nabla w_h)_{\mathcal{T}_h} - \langle \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} w_h \rangle_{\partial\mathcal{T}_h} - (\sigma y_h, w_h)_{\mathcal{T}_h}\\
&\qquad + (w_h, \boldsymbol{\beta}\cdot\nabla y_h)_{\mathcal{T}_h} - \langle \widehat{w}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} y_h \rangle_{\partial\mathcal{T}_h} + ((\sigma + \nabla\cdot\boldsymbol{\beta}) w_h, y_h)_{\mathcal{T}_h}\\
&\quad = -\langle \boldsymbol{\beta}\cdot\boldsymbol{n}(y_h - \widehat{y}_h^o), w_h - \widehat{w}_h^o \rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle y_h, \boldsymbol{\beta}\cdot\boldsymbol{n} w_h \rangle_{\partial\mathcal{T}_h} - \langle \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} w_h \rangle_{\partial\mathcal{T}_h} - \langle \widehat{w}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} y_h \rangle_{\partial\mathcal{T}_h}\\
&\quad = 0,
\end{aligned}$$

where we used $\langle \boldsymbol{\beta}\cdot\boldsymbol{n}, \widehat{y}_h^o \widehat{w}_h^o \rangle_{\partial\mathcal{T}_h} = 0$. This proves our result. ∎

3. Stability.

To perform the stability and error analysis for the convection dominated boundary control problem, we need to assume some conditions on the velocity vector field $\boldsymbol{\beta}$ and the effective reaction function $\bar{\sigma} := \sigma + \frac{1}{2}\nabla\cdot\boldsymbol{\beta}$.

(A1) $\bar{\sigma}$ has a nonnegative lower bound, i.e,

$$\sigma_0 := \inf_{\boldsymbol{x}\in\Omega} \bar{\sigma} \geq 0. \tag{3.1}$$

(A2) $\boldsymbol{\beta}$ has no closed curves and

$$|\boldsymbol{\beta}(\boldsymbol{x})| \neq 0 \ \forall \boldsymbol{x} \in \Omega.$$

(A3) $\varepsilon < \min_{T\in\mathcal{T}_h}\{h_T\}$.

We note that we have already assumed (A1) in (2.1) in Theorem 2.1. We repeat the assumption here to highlight it. Also, since we are interested in the convection dominated case, (A3) is a reasonable assumption. As shown in [3], assumption (A2) implies for any integer $k \geq 0$, there exists a function $\psi \in W^{k+1,\infty}(\Omega)$ such that for all $\boldsymbol{x} \in \Omega$, we have

$$\boldsymbol{\beta}\cdot\nabla\psi \geq 2\beta_0 > 0, \tag{3.2}$$

where $\beta_0 := \|\boldsymbol{\beta}\|_{0,\infty}/L$ and L is the diameter of Ω. We use assumption (A3) in the analysis to remove the assumption on the meshes. Specifically, in the proofs of Lemmas 4.6 and 4.10, we use assumption (A3) and a local inverse inequality to replace a global inverse inequality that has been used in all previous Dirichlet boundary control works. Therefore, we only assume $\{\mathcal{T}_h\}$ is a conforming simplex partition of Ω. All previous works on Dirichlet boundary control problems required a conforming quasi-uniform mesh. In the future, we hope to perform an a posteriori error analysis for the convection dominated boundary control problem.

Remark 3.1. If $\sigma_0 \geq 0$, then assumption (A2) is the minimal known requirement that can be used to establish stability and error analysis results for numerical methods; see, e.g., [3, 23]. If instead $\sigma_0 > 0$, then we don't need to assume (A2) and the numerical analysis is less technical. Specifically, we don't need to prove Theorem 3.14 below if $\sigma_0 > 0$.

3.1. Preliminary material. For any nonnegative integer j, we define the L^2-projections Π_j^o and Π_j^∂ as follows: for any $T \in \mathcal{T}_h$, $E \subset \partial T$, $v \in L^2(T)$, $q \in L^2(E)$, find $\Pi_j^o v \in \mathcal{P}_j(T)$ and $\Pi_j^\partial q \in \mathcal{P}_j(E)$ satisfying

$$(\Pi_j^o v, w_j)_T = (v, w_j)_T \quad \forall w_j \in \mathcal{P}_j(T), \tag{3.3a}$$

$$\langle \Pi_j^\partial q, r_j \rangle_E = \langle q, r_j \rangle_E \quad \forall r_j \in \mathcal{P}_j(E). \tag{3.3b}$$

We also define $\widetilde{\Pi}_k^\partial$ as

$$\widetilde{\Pi}_k^\partial|_E = \begin{cases} \Pi_k^\partial|_E, & E \in \mathcal{E}_h^o, \\ 0, & E \in \mathcal{E}_h^\partial. \end{cases}$$

Then $\widetilde{\Pi}_k^\partial$ is an operator mapping $L^2(\mathcal{E}_h)$ to M_h^o.

We first give an approximation property from [6, Theorem 4.3.8, Proposition 4.1.9], and then we prove the basic stability and approximation properties for L^2 projections.

LEMMA 3.2. *Let $m \geq 1$ be an integer. For any $T \in \mathcal{T}_h$, $v \in H^m(T)$, and integer s satisfying $0 \leq s \leq m$, there exists $I_{m-1}v \in \mathcal{P}_{m-1}(T)$ such that*

$$|v - I_{m-1}v|_{s,T} \leq C h_T^{m-s} |v|_{m,T}, \tag{3.4a}$$

$$\|v - I_{m-1}v\|_{0,\infty,T} \leq C h_T |v|_{1,\infty,T}. \tag{3.4b}$$

LEMMA 3.3. *Let s be a nonnegative integer. For any nonnegative integer j, let m be an integer satisfying $0 \leq m \leq j+1$ and let $\ell \in \{0,1\}$. For all $T \in \mathcal{T}_h$, $E \in \mathcal{E}_h$, it holds that*

$$|\Pi_\ell^o v|_{j,T} \leq C |v|_{j,T} \qquad \forall v \in H^j(T), \tag{3.5a}$$

$$\|\Pi_j^\partial v\|_E \leq \|v\|_E \qquad \forall v \in L^2(E), \tag{3.5b}$$

$$|v - \Pi_j^o v|_{s,T} \leq C h_T^{m-s} |v|_{m,T} \qquad \forall v \in H^m(T),\ 0 \leq s \leq m, \tag{3.5c}$$

$$|v - \Pi_0^o v|_{\ell,\infty,T} \leq C h_T^{1-\ell} |v|_{1,\infty,T} \qquad \forall v \in W^{1,\infty}(T), \tag{3.5d}$$

$$|v - \Pi_j^o v|_{\partial T} \leq C h_T^{m-1/2} |v|_{m,T} \qquad \forall v \in H^m(T), m \geq 1, \tag{3.5e}$$

$$\|w\|_{\partial T} \leq C h_T^{-1/2} \|w\|_T \qquad \forall w \in W_h. \tag{3.5f}$$

Proof. Equation (3.5a) follows from (3.5c); (3.5b) follows from the definition of L^2 projection; (3.5e) follows from (3.5c) and the trace inequality; and (3.5f) follows from the trace inequality and inverse inequality. The only thing left is to prove (3.5c) and (3.5d).

For (3.5c), in view of (3.4a), an inverse inequality, and the fact that $\|\Pi_j^o v\|_{0,T} \leq \|v\|_{0,T}$, for $1 \leq m \leq j+1$ we have

$$\begin{aligned}
|v - \Pi_j^o v|_{s,T} &\leq |v - I_{m-1}v|_{s,T} + |I_{m-1}v - \Pi_j^o v|_{s,T} \\
&= |v - I_{m-1}v|_{s,T} + |\Pi_j^o (I_{m-1}v - v)|_{s,T} \\
&\leq |v - I_{m-1}v|_{s,T} + Ch_T^{-s} \|\Pi_j^o (I_{m-1}v - v)\|_{0,T} \\
&\leq |v - I_{m-1}v|_{s,T} + Ch_T^{-s} \|I_{m-1}v - v\|_{0,T} \\
&\leq Ch_T^{m-s} |v|_{m,T}.
\end{aligned}$$

As for (3.5d), $\ell = 1$ is obvious and therefore we set $\ell = 0$. By a standard scaling argument, the following stability result holds:

$$\|\Pi_0^o v\|_{0,\infty,T} \leq C\|v\|_{0,\infty,T}. \tag{3.6}$$

By an inverse inequality, (3.6), and (3.4b) we get

$$\begin{aligned}
\|v - \Pi_0^o v\|_{0,\infty,T} &\leq \|v - I_0 v\|_{0,\infty,T} + \|I_0 v - \Pi_0^o v\|_{0,\infty,T} \\
&= \|v - I_0 v\|_{0,\infty,T} + \|\Pi_0^o (I_0 v - v)\|_{0,\infty,T} \\
&\leq C\|v - I_0 v\|_{0,\infty,T} \\
&\leq Ch_T |v|_{1,\infty,T}.
\end{aligned}$$

□

Next, we extend (3.5c) to real numbers s and m. We first recall the following classical results.

PROPOSITION 3.4 (see [6, Proposition 14.1.5]). *Given Banach spaces $A_1 \hookrightarrow A_0$ and $B_1 \hookrightarrow B_0$, let $\mathcal{K}$ be a bounded linear operator from A_i into B_i $(i = 0, 1)$. Then $\mathcal{K} : A_{\theta,p} \to B_{\theta,p}$ is a bounded linear operator for any $0 < \theta < 1$, $1 \leq p \leq \infty$. Moreover,*

$$\|\mathcal{K}\|_{A_{\theta,p} \to B_{\theta,p}} \leq \|\mathcal{K}\|_{A_0 \to B_0}^{1-\theta} \|\mathcal{K}\|_{A_1 \to B_1}^{\theta},$$

where $A_{\theta,p} := [A_0, A_1]_{\theta,p}$, $B_{\theta,p} := [B_0, B_1]_{\theta,p}$. See [6, p. 372] *for detailed definitions of $A_{\theta,p}$ and $B_{\theta,p}$.*

PROPOSITION 3.5 (cf. [6, Theorem 14.2.7]). *If Ω has a Lipschitz boundary, then*

$$[H^m(\Omega), H^\ell(\Omega)]_{\theta,2} = H^{(1-\theta)m + \theta\ell}(\Omega)$$

for all real numbers m and ℓ with $0 < \theta < 1$.

Now will give results similar to (3.5c).

LEMMA 3.6. *Let s be a real number. For any nonnegative integer j, let m be a real number satisfying $0 \leq m \leq j+1$. For all $T \in \mathcal{T}_h$, it holds that*

$$\|v - \Pi_j^o v\|_{s,T} \leq Ch_T^{m-s} \|v\|_{m,T} \qquad \forall v \in H^m(T),\ 0 \leq s \leq m, \tag{3.7a}$$

$$\|v - \Pi_j^o v\|_{\partial T} \leq Ch_T^{m-1/2} \|v\|_{m,T} \qquad \forall v \in H^m(T), m > 1/2. \tag{3.7b}$$

Proof. We split the proof of (3.7a) into two steps.

(1) Assume m is an integer and $s \in (0, m)$. We take $A_0 = A_1 = H^m(\Omega)$, $B_0 = H^0(\Omega)$, $B_1 = H^m(\Omega)$, and $\theta = s/m$. Then, by Proposition 3.5, we have

$$\begin{aligned}
A_{\theta,2} &= [A_0, A_1]_{\theta,2} = [H^m(\Omega), H^m(\Omega)]_{\theta,2} = H^m(\Omega), \\
B_{\theta,2} &= [B_0, B_1]_{\theta,2} = [H^0(\Omega), H^m(\Omega)]_{\theta,2} = H^s(\Omega).
\end{aligned} \tag{3.8}$$

Hence, we get

$$
\begin{aligned}
&\frac{\|(Id-\Pi_j^o)v\|_{s,T}}{\|v\|_{m,T}}\\
&\quad\le \|(Id-\Pi_j^o)\|_{H^m(T)\to H^s(T)}\\
&\quad= \|(Id-\Pi_j^o)\|_{A_{\theta,2}\to B_{\theta,2}} && \text{by (3.8)}\\
&\quad\le \|(Id-\Pi_j^o)\|_{A_0\to B_0}^{1-\theta}\|(Id-\Pi_j^o)\|_{A_1\to B_1}^{\theta} && \text{by Proposition 3.4}\\
&\quad= \|(Id-\Pi_j^o)\|_{H^m(T)\to H^0(T)}^{1-\theta}\|(Id-\Pi_j^o)\|_{H^m(T)\to H^m(T)}^{\theta} && \text{by (3.8)}\\
&\quad\le Ch_T^{m(1-\theta)} && \text{by (3.5c)}\\
&\quad= Ch_T^{m-s}.
\end{aligned}
$$

By (3.5c), the above inequality holds for $s=0$ and $s=m$. This implies for $s\in[0,m]$ we have

$$\|v-\Pi_j^o v\|_{s,T}\le Ch_T^{m-s}\|v\|_{m,T}. \tag{3.9}$$

(2) Assume m is not an integer and $s\in[0,m]$. Let $[m]$ be the largest integer which is no larger than m, so that $s\in[0,[m]+1]$. We take $A_0=H^s(\Omega)$, $A_1=H^{[m]+1}(\Omega)$, $B_0=B_1=H^s(\Omega)$, $\theta=\frac{m-s}{[m]+1-s}\in(0,1)$. Remembering that the inequality (3.9) holds for all integer m, hence we have

$$\|v-\Pi_j^o v\|_{s,T}\le Ch_T^{[m]+1-s}\|v\|_{[m]+1,T}. \tag{3.10}$$

$$
\begin{aligned}
\frac{\|(Id-\Pi_j^o)v\|_{s,T}}{\|v\|_{m,T}} &\le \|(Id-\Pi_j^o)\|_{H^m(T)\to H^s(T)}\\
&\le \|(Id-\Pi_j^o)\|_{H^s(T)\to H^s(T)}^{1-\theta}\|(Id-\Pi_j^o)\|_{H^{[m]+1}(T)\to H^s(T)}^{\theta}\\
&\le Ch_T^{([m]+1-s)\theta} && \text{by (3.10)}\\
&= Ch_T^{m-s}.
\end{aligned}
$$

Hence, (1) and (2) prove that (3.7a) holds for any real number $m\in[0,j+1]$ and real number $s\in[0,m]$.

The result (3.7b) follows by (3.7a) and [6, Lemma 1.6.6] for $m\ge 1$ and [22, Lemma 7.2] for $m\in(1/2,1)$. ∎

In addition, we have the following superapproximation results; a similar result can be found in [3].

LEMMA 3.7. *Let $T\in\mathcal{T}_h, E\subset\partial T$. Then, for any $u_h\in\mathcal{P}_k(T)$ and $\eta\in W^{1,\infty}(T)$, there holds*

$$\|\eta u_h-\Pi_k^o(\eta u_h)\|_T\le Ch_T|\eta|_{1,\infty,T}\|u_h\|_T, \tag{3.11a}$$

$$|\eta u_h-\Pi_k^o(\eta u_h)|_{1,T}\le C|\eta|_{1,\infty,T}\|u_h\|_T, \tag{3.11b}$$

$$\|\eta u_h-\Pi_k^o(\eta u_h)\|_{\partial T}\le Ch_T^{1/2}|\eta|_{1,\infty,T}\|u_h\|_T, \tag{3.11c}$$

$$\|\eta u_h-\Pi_k^\partial(\eta u_h)\|_E\le Ch_T|\eta|_{1,\infty,T}\|u_h\|_E. \tag{3.11d}$$

Proof. We notice that (3.11c) follows from (3.11a), (3.11b), and the trace inequality. Next, for $j \in \{0, 1\}$, we have

$$\begin{aligned}
|\eta u_h - \Pi_k^o(\eta u_h)|_{j,T} &\le |\eta u_h - (\Pi_0^o\eta)u_h|_{j,T} + |(\Pi_0^o\eta)u_h - \Pi_k^o(\eta u_h)|_{j,T} \\
&= |(\eta - \Pi_0^o\eta)u_h|_{j,T} + |\Pi_k^o((\Pi_0^o\eta)u_h - \eta u_h)|_{j,T} \\
&\le C|(\eta - \Pi_0^o\eta)u_h|_{j,T} && \text{by (3.5a)} \\
&\le C|\eta - \Pi_0^o\eta|_{j,\infty}\|u_h\|_T + \|\eta - \Pi_0^o\eta\|_{0,\infty}|u_h|_{j,T} \\
&\le Ch_T^{1-j}|\eta|_{1,\infty,T}\|u_h\|_T && \text{by (3.5d).}
\end{aligned}$$

This proves (3.11a) and (3.11b). Similarly, for $E \subset \partial T$, we have

$$\begin{aligned}
\|\eta u_h - \Pi_k^\partial(\eta u_h)\|_E &\le \|\eta u_h - (\Pi_0^o\eta)u_h\|_E + \|(\Pi_0^o\eta)u_h - \Pi_k^\partial(\eta u_h)\|_E \\
&= \|(\eta - \Pi_0^o\eta)u_h\|_E + \|\Pi_k^\partial((\Pi_0^o\eta)u_h - \eta u_h)\|_E \\
&\le C\|(\eta - \Pi_0^o\eta)u_h\|_E && \text{by (3.5b)} \\
&\le C\|\eta - \Pi_0^o\eta\|_{0,\infty,T}\|u_h\|_E \\
&\le Ch_T|\eta|_{1,\infty,T}\|u_h\|_E && \text{by (3.5d).}
\end{aligned}$$

This proves (3.11d). ∎

For the analysis of the low regularity case, we need the following result from [38].

LEMMA 3.8. *If $\ell \ge 1$ is an integer that is large enough, then there exists an interpolation operator $\mathcal{I}_h^c : W_h \times M_h^o \to H_0^1(\Omega) \cap \mathcal{P}_{k+\ell}^{\mathcal{T}_h}$ such that for all $(w_h, \widehat{w}_h^o, v_h, \widehat{v}_h^o) \in [W_h \times M_h^o]^2$, for all $T \in \mathcal{T}_h$, and for all $E \in \mathcal{E}_h$, we have*

$$(\mathcal{I}_h^c(w_h, \widehat{w}_h^o), v_h)_T = (w_h, v_h)_T, \tag{3.12a}$$

$$\langle \mathcal{I}_h^c(w_h, \widehat{w}_h^o), \widehat{v}_h^o\rangle_E = \langle \widehat{w}_h^o, \widehat{v}_h^o\rangle_E, \tag{3.12b}$$

$$\|\nabla \mathcal{I}_h^c(w_h, \widehat{w}_h^o)\|_{\mathcal{T}_h} \le C\left(\|\nabla w_h\|_{\mathcal{T}_h} + \|h_E^{-1/2}(w_h - \widehat{w}_h^o)\|_{\partial\mathcal{T}_h}\right), \tag{3.12c}$$

$$\|w_h - \mathcal{I}_h^c(w_h, \widehat{w}_h^o)\|_{\mathcal{T}_h} \le Ch\left(\|\nabla w_h\|_{\mathcal{T}_h} + \|h_E^{-1/2}(w_h - \widehat{w}_h^o)\|_{\partial\mathcal{T}_h}\right), \tag{3.12d}$$

where $\mathcal{P}_{k+\ell}^{\mathcal{T}_h} = \{w_h \in L^2(\Omega) : w_h|_T \in \mathcal{P}_{k+\ell}(T) \text{ for all } T \in \mathcal{T}_h\}$.

3.2. Proof of the stability of (2.6). Next, we present the stability of the above HDG method for the convection dominated Dirichlet boundary control problem. We follow a similar strategy to [3, 23]. We first collect some basic equalities and inequalities, which are used frequently in our paper.

LEMMA 3.9. *For all $(w_h, \widehat{w}_h^o) \in W_h \times M_h^o$, we have*

$$(w_h, \boldsymbol{\beta}\cdot\nabla w_h)_{\mathcal{T}_h} = \frac{1}{2}\langle \boldsymbol{\beta}\cdot\boldsymbol{n} w_h, w_h\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}(\nabla\cdot\boldsymbol{\beta} w_h, w_h)_{\mathcal{T}_h}, \tag{3.13a}$$

$$\frac{1}{2}\langle \boldsymbol{\beta}\cdot\boldsymbol{n} w_h, w_h\rangle_{\partial\mathcal{T}_h} - \langle \widehat{w}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} w_h\rangle_{\partial\mathcal{T}_h} = \frac{1}{2}\langle \boldsymbol{\beta}\cdot\boldsymbol{n}(w_h - \widehat{w}_h^o), w_h - \widehat{w}_h^o\rangle_{\partial\mathcal{T}_h}, \tag{3.13b}$$

$$\|w_h\|_{\mathcal{T}_h}^2 \le C\|\nabla w_h\|_{\mathcal{T}_h}^2 + C\sum_{T\in\mathcal{T}_h} h_T^{-1}\|w_h - \widehat{w}_h^o\|_{\partial T}^2. \tag{3.13c}$$

The identity (3.13a) can be obtained by integration by parts and the proof of (3.13b) follows from the fact $\langle \boldsymbol{\beta}\cdot\boldsymbol{n}\widehat{w}_h^o, \widehat{w}_h^o\rangle_{\partial\mathcal{T}_h} = 0$. For the last inequality (3.13c), we refer to [42, p. 354] for the proof.

Next, we define some seminorms.

DEFINITION 3.10. *For all* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o) \in \boldsymbol{V}_h \times W_h \times M_h^o$, *define*

$$\|(y_h, \widehat{y}_h^o)\|_{W,w}^2 := \varepsilon\|\nabla y_h\|_{\mathcal{T}_h}^2 + \|\tau^{1/2}(y_h - \widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2 + \|\bar{\sigma}^{1/2} y_h\|_{\mathcal{T}_h}^2, \tag{3.14a}$$

$$\|(y_h, \widehat{y}_h^o)\|_{W}^2 := \varepsilon\|\nabla y_h\|_{\mathcal{T}_h}^2 + \|\tau^{1/2}(y_h - \widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2 + \|(\beta_0+\bar{\sigma})^{1/2} y_h\|_{\mathcal{T}_h}^2, \tag{3.14b}$$

$$\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_{w}^2 := \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \|(y_h, \widehat{y}_h^o)\|_{W,w}^2, \tag{3.14c}$$

$$\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|^2 := \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \|(y_h, \widehat{y}_h^o)\|_{W}^2 . \tag{3.14d}$$

It is easy to see that the seminorm $\|(\cdot,\cdot)\|_W$ is a norm since $\beta_0 > 0$, and hence $\|(\cdot,\cdot,\cdot)\|$ is also a norm. To prove the seminorms $\|(\cdot,\cdot)\|_{W,w}$ and $\|(\cdot,\cdot,\cdot)\|_w$ are norms, we just need to show $\|(\cdot,\cdot)\|_{W,w}$ is a norm.

LEMMA 3.11. $\|(\cdot,\cdot)\|_{W,w}$ *is a norm for the space* $W_h \times M_h^o$.

Proof. It is obvious that we only need to show that $\|(y_h, \widehat{y}_h^o)\|_{W,w} = 0$ implies $\widehat{y}_h^o|_\Gamma = y_h = 0$. This is true because y_h is piecewise constant on $\mathcal{T}_h$ and $y_h = \widehat{y}_h^o$ on $\mathcal{E}_h$; therefore, $y_h = \widehat{y}_h^o$ are constants. Since $\widehat{y}_h^o|_\Gamma = 0$, we have $y_h = \widehat{y}_h^o = 0$. ∎

LEMMA 3.12 (stability in weak norm). *For all* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o) \in \boldsymbol{V}_h \times W_h \times M_h^o$, *the following stability results hold:*

$$\sup_{0\neq(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\in\boldsymbol{V}_h\times W_h\times M_h^o} \frac{\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o)}{\|(\boldsymbol{r}_h, w_h, \widehat{w}_h^o)\|_w} \geq C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w\,, \tag{3.15a}$$

$$\sup_{0\neq(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\in\boldsymbol{V}_h\times W_h\times M_h^o} \frac{\mathcal{B}_2(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o)}{\|(\boldsymbol{r}_h, w_h, \widehat{w}_h^o)\|_w} \geq C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w\,. \tag{3.15b}$$

Proof. We only prove the first inequality; the second can be obtained by the same argument. First, let $(\boldsymbol{r}_h, w_h, \widehat{w}_h^o) = (q_h, -y_h, -\widehat{y}_h^o)$ in the definition of $\mathcal{B}_1$ in (2.4) to get

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{q}_h, -y_h, -\widehat{y}_h^o)\\
&\quad= \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \langle\tau_1(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad - (y_h, \boldsymbol{\beta}\nabla y_h)_{\mathcal{T}_h} + \langle\widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} y_h\rangle_{\partial\mathcal{T}_h} + (\sigma y_h, y_h)_{\mathcal{T}_h}\\
&\quad= \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \langle\tau_1(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad - \frac{1}{2}\langle\boldsymbol{\beta}\cdot\boldsymbol{n} y_h, y_h\rangle_{\partial\mathcal{T}_h} + \langle\widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} y_h\rangle_{\partial\mathcal{T}_h} + ((\sigma + \frac{1}{2}\nabla\cdot\boldsymbol{\beta})y_h, y_h)_{\mathcal{T}_h} \qquad \text{by (3.13a)}\\
&\quad= \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \langle\tau_1(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad - \frac{1}{2}\langle\boldsymbol{\beta}\cdot\boldsymbol{n}(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h} + ((\sigma + \frac{1}{2}\nabla\cdot\boldsymbol{\beta})y_h, y_h)_{\mathcal{T}_h} \qquad \text{by (3.13b)}\\
&\quad= \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\tau}(y_h - \widehat{y}_h^o)\right\|_{\partial\mathcal{T}_h}^2 + \|\bar{\sigma}^{1/2} y_h\|_{\mathcal{T}_h}^2 \qquad \text{by (3.1).}
\end{aligned} \tag{3.16}$$

Next, let $(\boldsymbol{r}_h, w_h, \widehat{w}_h^o) = (\varepsilon\nabla y_h, 0, 0)$ to get

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \varepsilon\nabla y_h, 0, 0)\\
&\quad= (\boldsymbol{q}_h, \nabla y_h)_{\mathcal{T}_h} - \varepsilon(y_h, \nabla\cdot\nabla y_h)_{\mathcal{T}_h} + \varepsilon\langle\widehat{y}_h^o, \nabla y_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\quad= \varepsilon\|\nabla y_h\|_{\mathcal{T}_h}^2 + \varepsilon\langle\widehat{y}_h^o - y_h, \nabla y_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{q}_h, \nabla y_h)_{\mathcal{T}_h}\\
&\quad\geq \varepsilon\|\nabla y_h\|_{\mathcal{T}_h}^2 - \varepsilon^{1/2}h^{-1/2}\|\widehat{y}_h^o - y_h\|_{\partial\mathcal{T}_h}\varepsilon^{1/2}\|\nabla y_h\|_{\mathcal{T}_h} - \|\boldsymbol{q}_h\|_{\mathcal{T}_h}\|\nabla y_h\|_{\mathcal{T}_h}\\
&\quad\geq \frac{\varepsilon}{2}\|\nabla y_h\|_{\mathcal{T}_h}^2 - C_0(\|\tau^{1/2}(\widehat{y}_h^o - y_h)\|_{\partial\mathcal{T}_h}^2 + \varepsilon^{-1}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2),
\end{aligned} \tag{3.17}$$

where C_0 is a fixed positive constant. The definitions of $\|(\cdot,\cdot,\cdot)\|$ in (3.14d) and $\|(\cdot,\cdot,\cdot)\|_w$ in (3.14c) imply

$$\|(\varepsilon\nabla y_h,0,0)\| = \|(\varepsilon\nabla y_h,0,0)\|_w \le C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|_w\,. \tag{3.18}$$

Finally, we take $(\boldsymbol{r}_h,w_h,\widehat{w}_h^o) = (\frac{1}{2}+C_0)(\boldsymbol{q}_h,-y_h,-\widehat{y}_h^o) + (\varepsilon\nabla y_h,0,0)$ to obtain

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h,y_h,\widehat{y}_h^o;\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\\
&\quad\ge \frac{1}{2}\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|_w^2 && \text{by (3.16) and (3.17)}\\
&\quad\ge C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|_w\,\|(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\|_w && \text{by (3.18).}
\end{aligned}\tag{3.19}$$

This completes our proof. ◻

For later use, by (3.18), for any $(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\in \boldsymbol{V}_h\times W_h\times M_h^o$, we have

$$\|(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\| \le C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\| + \|(\varepsilon\nabla y_h,0,0)\| \le C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|\,. \tag{3.20}$$

Remark 3.13. The existence of a unique solution to the HDG discretization (2.6) of the optimality system now follows similarly to [32]; we omit the details. Also, to obtain the L^2 error estimates for the state y_h, Lemma 3.12 is not sufficient since the effective reaction term $\bar{\sigma}^{1/2}$ can equal zero at some points; therefore, it is possible for the term $\|\bar{\sigma}^{1/2}y_h\|_{\mathcal{T}_h}$ in the definition of $\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|_w$ to equal zero for some y_h. Therefore, we need a refined analysis technique to derive a strong stability result that contains the norm $\|y_h\|_{\mathcal{T}_h}$.

THEOREM 3.14 (stability in strong norm). *If assumptions* (A1) *and* (A2) *hold, then there exists h_0, independent of ε, such that the following stability results hold: for all $(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\in \boldsymbol{V}_h\times W_h\times M_h^o$ with $h\le h_0$,*

$$\sup_{0\neq(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\in \boldsymbol{V}_h\times W_h\times M_h^o} \frac{\mathcal{B}_1(\boldsymbol{q}_h,y_h,\widehat{y}_h^o;\boldsymbol{r}_h,w_h,\widehat{w}_h^o)}{\|(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\|} \ge C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|\,, \tag{3.21a}$$

$$\sup_{0\neq(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\in \boldsymbol{V}_h\times W_h\times M_h^o} \frac{\mathcal{B}_2(\boldsymbol{q}_h,y_h,\widehat{y}_h^o;\boldsymbol{r}_h,w_h,\widehat{w}_h^o)}{\|(\boldsymbol{r}_h,w_h,\widehat{w}_h^o)\|} \ge C\,\|(\boldsymbol{q}_h,y_h,\widehat{y}_h^o)\|\,. \tag{3.21b}$$

Proof. We only prove (3.21a), and we split the proofs into two steps.

Step 1. Let $\psi\in W^{1,\infty}(\Omega)$ satisfy (3.2). We take

$$(\boldsymbol{r}_h,w_h,\widehat{w}_h^o) = (\boldsymbol{r}_1,w_1,\widehat{w}_1) = (\boldsymbol{0},-e^{-\psi}y_h,-e^{-\psi}\widehat{y}_h^o) \tag{3.22}$$

in the definition of $\mathcal{B}_1$ in (2.4) to obtain

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h,y_h,\widehat{y}_h^o;\boldsymbol{r}_1,w_1,\widehat{w}_1)\\
&\quad= [(e^{-\psi}y_h,\nabla\cdot\boldsymbol{q}_h)_{\mathcal{T}_h} - \langle e^{-\psi}\widehat{y}_h^o,\boldsymbol{q}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}] + \langle\tau_1(y_h-\widehat{y}_h^o),e^{-\psi}y_h-e^{-\psi}\widehat{y}_h^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad+[-(y_h,\boldsymbol{\beta}\cdot\nabla(e^{-\psi}y_h))_{\mathcal{T}_h} + \langle\widehat{y}_h^o,\boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}y_h\rangle_{\partial\mathcal{T}_h}] + (\sigma y_h,e^{-\psi}y_h)_{\mathcal{T}_h}\\
&\quad= S_1+S_2+S_3+S_4.
\end{aligned}$$

Next, we estimate $\{S_i\}_{i=1}^4$ term by term. First,

$$
\begin{aligned}
S_1 &= -(\nabla(e^{-\psi} y_h), \boldsymbol{q}_h)_{\mathcal{T}_h} - \langle e^{-\psi}(\widehat{y}_h^o - y_h), \boldsymbol{q}_h \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&= -(y_h \nabla e^{-\psi} + e^{-\psi}\nabla y_h, \boldsymbol{q}_h)_{\mathcal{T}_h} - \langle e^{-\psi}(\widehat{y}_h^o - y_h), \boldsymbol{q}_h \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\le C(\varepsilon^{1/2}\|y_h\|_{\mathcal{T}_h} + \varepsilon^{1/2}\|\nabla y_h\|_{\mathcal{T}_h} + \|\tau^{1/2}(\widehat{y}_h^o - y_h)\|_{\partial\mathcal{T}_h})\varepsilon^{-1/2}\|\boldsymbol{q}_h\|_{\mathcal{T}_h}\\
&\le C(\varepsilon^{1/2}\|\nabla y_h\|_{\mathcal{T}_h} + \|\tau^{1/2}(\widehat{y}_h^o - y_h)\|_{\partial\mathcal{T}_h})\varepsilon^{-1/2}\|\boldsymbol{q}_h\|_{\mathcal{T}_h} && \text{by (3.13c)}\\
&\le C\,\|(y_h, \widehat{y}_h^o)\|_{W,w}\,\varepsilon^{-1/2}\|\boldsymbol{q}_h\|_{\mathcal{T}_h} && \text{by (3.14a)}\\
&\le C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w^2 && \text{by (3.14c).}
\end{aligned}
$$

Second, to estimate the term S_3, let $\varphi = e^{-\psi}$ in (3.13a) to obtain

$$
\begin{aligned}
-(y_h, \boldsymbol{\beta}\cdot\nabla(e^{-\psi}y_h))_{\mathcal{T}_h} &= \frac{1}{2}(\nabla\cdot\boldsymbol{\beta} e^{-\psi}y_h, y_h)_{\mathcal{T}_h} + \frac{1}{2}(y_h, e^{-\psi}y_h\boldsymbol{\beta}\cdot\nabla\psi)_{\mathcal{T}_h}\\
&\quad - \frac{1}{2}\langle y_h, \boldsymbol{\beta}\cdot\boldsymbol{n} e^{-\psi} y_h\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
S_3 &= -(y_h, \boldsymbol{\beta}\cdot\nabla(e^{-\psi}y_h))_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}y_h\rangle_{\partial\mathcal{T}_h}\\
&= \frac{1}{2}(\nabla\cdot\boldsymbol{\beta}e^{-\psi}y_h, y_h)_{\mathcal{T}_h} + \frac{1}{2}(y_h, e^{-\psi}y_h\boldsymbol{\beta}\cdot\nabla\psi)_{\mathcal{T}_h}\\
&\quad - \frac{1}{2}\langle y_h, \boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}y_h\rangle_{\partial\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}y_h\rangle_{\partial\mathcal{T}_h}\\
&= \frac{1}{2}(\nabla\cdot\boldsymbol{\beta}e^{-\psi}y_h, y_h)_{\mathcal{T}_h} + \frac{1}{2}(e^{-\psi}\boldsymbol{\beta}\cdot\nabla\psi y_h, y_h)_{\mathcal{T}_h}\\
&\quad - \frac{1}{2}\langle \boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h} && \text{by (3.13b)}\\
&\ge \frac{1}{2}(\nabla\cdot\boldsymbol{\beta}e^{-\psi}y_h, y_h)_{\mathcal{T}_h} + (\beta_0 e^{-\psi}y_h, y_h)_{\mathcal{T}_h} && \text{by (3.2)}\\
&\quad - \frac{1}{2}\langle \boldsymbol{\beta}\cdot\boldsymbol{n}e^{-\psi}(y_h - \widehat{y}_h^o), y_h - \widehat{y}_h^o\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Therefore, by the definition of τ in (2.3h), there exists a positive constant C_0 such that

$$
\begin{aligned}
S_2 + S_3 + S_4 &\ge \|e^{-\psi/2}(\beta_0 + \bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2 + \|e^{-\psi/2}\tau^{1/2}(y_h - \widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2\\
&\ge C_0\|(\beta_0 + \bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2.
\end{aligned}
$$

This implies that there exist positive constants C_1 and C_2 such that

$$
\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_1, w_1, \widehat{w}_1) \ge C_1\|(\beta_0 + \bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2 - C_2\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w^2. \tag{3.23}
$$

Moreover, we have

$$
\begin{aligned}
\|(\boldsymbol{r}_1, w_1, \widehat{w}_1)\|^2 &= \big\|(\boldsymbol{0}, -e^{-\psi}y_h, -e^{-\psi}\widehat{y}_h^o)\big\| && \text{by (3.22)}\\
&= \varepsilon\|\nabla(e^{-\psi}y_h)\|_{\mathcal{T}_h}^2 + \|\tau^{1/2}e^{-\psi}(y_h - \widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2\\
&\quad + \|(\beta_0 + \bar{\sigma})^{1/2}e^{-\psi}y_h\|_{\mathcal{T}_h}^2 && \text{by (3.14d)}\\
&\le C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|^2 && \text{by (3.13c),} \qquad (3.24)
\end{aligned}
$$

where we used $\nabla(e^{-\psi}y_h) = y_h\nabla e^{-\psi} + e^{-\psi}\nabla y_h$ in (3.24).

Step 2. Let $R_k^o = \mathbb{I} - \Pi_k^o$ and $\widetilde{R}_k^\partial = \mathbb{I} - \widetilde{\Pi}_k^\partial$, where $\mathbb{I}$ is the identity operator. We take

$$(\boldsymbol{r}_h, w_h, \widehat{w}_h^o) = (\boldsymbol{r}_2, w_2, \widehat{w}_2) = (\boldsymbol{0}, R_k^o(e^{-\psi} y_h), \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)) \tag{3.25}$$

in the definition of $\mathcal{B}_1$ and use the orthogonality properties of Π_k^o and $\widetilde{\Pi}_k^\partial$, integration by parts, and $\tau_1 = \tau_2 + \boldsymbol{\beta}\cdot\boldsymbol{n}$ to get

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_2, w_2, \widehat{w}_2)\\
&\quad= -(R_k^o(e^{-\psi} y_h), \nabla\cdot\boldsymbol{q}_h)_{\mathcal{T}_h} + \langle \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o), \boldsymbol{q}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\qquad - \langle \tau_1(y_h - \widehat{y}_h^o), R_k^o(e^{-\psi} y_h) - \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h}\\
&\qquad + (y_h, \boldsymbol{\beta}\cdot\nabla R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h} - \langle \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n} R_k^o(e^{-\psi} y_h)\rangle_{\partial\mathcal{T}_h}\\
&\qquad - (\sigma y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}.
\end{aligned} \tag{3.26}$$

The definitions of Π_k^o and $\widetilde{\Pi}_k^\partial$ in (3.3) imply

$$(R_k^o(e^{-\psi} y_h), \nabla\cdot\boldsymbol{q}_h)_{\mathcal{T}_h} = 0 \quad \text{and} \quad \langle \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o), \boldsymbol{q}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0.$$

Next, integration by parts gives

$$\begin{aligned}
&(y_h, \boldsymbol{\beta}\cdot\nabla R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}\\
&\quad= \langle \boldsymbol{\beta}\cdot\boldsymbol{n} y_h, R_k^o(e^{-\psi} y_h)\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{\beta}\nabla y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h} - (\nabla\cdot\boldsymbol{\beta} y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}\\
&\quad= \langle \boldsymbol{\beta}\cdot\boldsymbol{n}(y_h - \widehat{y}_h^o), R_k^o(e^{-\psi} y_h)\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{\beta}\cdot\boldsymbol{n}\widehat{y}_h^o, R_k^o(e^{-\psi} y_h)\rangle_{\partial\mathcal{T}_h}\\
&\qquad - (\boldsymbol{\beta}\cdot\nabla y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h} - (\nabla\cdot\boldsymbol{\beta} y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}\\
&\quad= \langle \boldsymbol{\beta}\cdot\boldsymbol{n}(y_h - \widehat{y}_h^o), R_k^o(e^{-\psi} y_h) - \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{\beta}\cdot\boldsymbol{n}(y_h - \widehat{y}_h^o), \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle \boldsymbol{\beta}\cdot\boldsymbol{n}\widehat{y}_h^o, R_k^o(e^{-\psi} y_h)\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{\beta}\cdot\nabla y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h} - (\nabla\cdot\boldsymbol{\beta} y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}.
\end{aligned}$$

Using $\tau_1 = \tau_2 + \boldsymbol{\beta}\cdot\boldsymbol{n}$ along with (3.26) and the above equalities gives

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_2, w_2, \widehat{w}_2)\\
&\quad= -\langle \tau_2(y_h - \widehat{y}_h^o), R_k^o(e^{-\psi} y_h) - \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle y_h - \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n}\widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{\beta}\cdot\nabla y_h, R_k^o(e^{-\psi} y_h))_{\mathcal{T}_h}\\
&\qquad - ((\sigma + \nabla\cdot\boldsymbol{\beta}) y_h, R_k^o e^{-\psi} y_h)_{\mathcal{T}_h}\\
&\quad= T_1 + T_2 + T_3 + T_4.
\end{aligned} \tag{3.27}$$

Before we estimate $\{T_i\}_{i=1}^4$, we first define $R_k^\partial = \mathbb{I} - \Pi_k^\partial$ and estimate the following term:

$$\begin{aligned}
&\|(R_k^o(e^{-\psi} y_h) - R_k^\partial(e^{-\psi}\widehat{y}_h^o))\|_{\partial\mathcal{T}_h}^2\\
&\quad\le C\sum_{T\in\mathcal{T}_h}\left(\|R_k^o(e^{-\psi} y_h)\|_{\partial T}^2 + \|R_k^\partial(e^{-\psi}\widehat{y}_h^o)\|_{\partial T}^2\right)\\
&\quad\le C\sum_{T\in\mathcal{T}_h}\left(h_T\|y_h\|_T^2 + h_T^2\|y_h - \widehat{y}_h^o\|_{\partial T}^2\right) \qquad\qquad \text{by (3.11c)–(3.11d),}\\
&\quad\le Ch^2\|y_h - \widehat{y}_h^o\|_{\partial\mathcal{T}_h}^2 + Ch\|y_h\|_{\mathcal{T}_h}^2.
\end{aligned} \tag{3.28}$$

Therefore,

$$|T_1| \le Ch^2\|\tau^{1/2}(y_h - \widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2 + Ch\|(\beta_0+\bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2. \tag{3.29}$$

Next, by the definition of R_k^o, we have

$$\begin{aligned}
T_2+T_3+T_4 &= -(\boldsymbol{\beta}\cdot\nabla y_h, R_k^o(e^{-\psi}y_h))_{\mathcal{T}_h} + \langle y_h - \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n}\widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h}\\
&\quad - ((\sigma+\nabla\cdot\boldsymbol{\beta})y_h, R_k^o(e^{-\psi}y_h))_{\mathcal{T}_h}\\
&= -(R_0^o(\boldsymbol{\beta})\cdot\nabla y_h, R_k^o(e^{-\psi}y_h))_{\mathcal{T}_h} + \langle y_h - \widehat{y}_h^o, \boldsymbol{\beta}\cdot\boldsymbol{n}\widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o)\rangle_{\partial\mathcal{T}_h}\\
&\quad + ((\sigma+\nabla\cdot\boldsymbol{\beta})y_h, R_k^o(e^{-\psi}y_h))_{\mathcal{T}_h}.
\end{aligned}$$

Hence,

$$\begin{aligned}
T_2+T_3+T_4 &\le C\left(\sum_{T\in\mathcal{T}_h} h_T^2\|\nabla y_h\|_T^2\right)^{1/2} h\|y_h\|_{\mathcal{T}_h} && \text{by (3.11a)}\\
&\quad + C\left(\sum_{T\in\mathcal{T}_h}\|y_h-\widehat{y}_h^o\|_{\partial T}^2\right)^{1/2}\left(\sum_{T\in\mathcal{T}_h} h_T^2\|y_h\|_{\partial T}^2\right)^{1/2} && \text{by (3.11d)}\\
&\quad + Ch(\|\boldsymbol{\beta}\|_{1,\infty} + \|\bar{\sigma}\|_{0,\infty})\|y_h\|_{\mathcal{T}_h}^2 && \text{by (3.11a)}\\
&\le Ch\|(\beta_0+\bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2 + \|\tau^{1/2}(y_h-\widehat{y}_h^o)\|_{\partial\mathcal{T}_h}^2. && (3.30)
\end{aligned}$$

From (3.27), (3.29), and (3.30), we get

$$\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_2, w_2, \widehat{w}_2^o) \ge -C_3h\|(\beta_0+\bar{\sigma})^{1/2}y_h\|_{\mathcal{T}_h}^2 - C_4\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w^2. \tag{3.31}$$

Using (3.28), we have

$$\begin{aligned}
\|(\boldsymbol{r}_2, w_2, \widehat{w}_2^o)\|^2 &= \left\|(\boldsymbol{0}, R_k^o(e^{-\psi}y_h), \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o))\right\|^2\\
&= \varepsilon\|\nabla R_k^o(e^{-\psi}y_h)\|_{\mathcal{T}_h}^2 + \|(\beta_0+\bar{\sigma})^{1/2}e^{-\psi}y_h\|_{\mathcal{T}_h}^2\\
&\quad + \|\tau^{1/2}(R_k^o(e^{-\psi}y_h) - \widetilde{R}_k^\partial(e^{-\psi}\widehat{y}_h^o))\|_{\partial\mathcal{T}_h}^2\\
&\le C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|^2 \qquad \text{by (3.13c).} && (3.32)
\end{aligned}$$

By (3.20), there exists a $(\boldsymbol{r}_0, w_0, \widehat{w}_0^o) \in \boldsymbol{V}_h\times W_h\times M_h^o$ such that

$$\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_0, w_0, \widehat{w}_0^o) \ge \|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|_w^2, \tag{3.33a}$$

$$\|(\boldsymbol{r}_0, w_0, \widehat{w}_0^o)\| \le C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|. \tag{3.33b}$$

Take h small enough so that we have $C_3h \le C_1/2$. Set $C_* = C_1 + C_2 + C_4$ and

$$(\boldsymbol{r}_h, w_h, \widehat{w}_h^o) = C_*(\boldsymbol{r}_0, w_0, \widehat{w}_0^o) + (\boldsymbol{r}_1, w_1, \widehat{w}_1^o) + (\boldsymbol{r}_2, w_2, \widehat{w}_2^o). \tag{3.34}$$

By (3.23), (3.31), and (3.33), we get

$$\mathcal{B}_1(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_h, w_h, \widehat{w}_h^o) \ge C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|^2 \ge C\,\|(\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\|\cdot\|(\boldsymbol{r}_h, w_h, \widehat{w}_h^o)\|,$$

which implies (3.21a). ∎

4. Error analysis. Next, we perform a convergence analysis for the convection dominated Dirichlet boundary control problem.

4.1. Assumptions and main result. Throughout, we assume Ω is a bounded convex polygonal domain in two dimensions. Therefore, the largest interior angle ω satisfies $\pi/3 \leq \omega < \pi$. Moreover, we assume the velocity vector field $\boldsymbol{\beta}$ and σ satisfy

$$\boldsymbol{\beta} \in [C(\overline{\Omega})]^2, \nabla \cdot \boldsymbol{\beta} \in L^\infty(\Omega), \sigma + \frac{1}{2}\nabla \cdot \boldsymbol{\beta} \geq 0, \nabla\nabla \cdot \boldsymbol{\beta} \in [L^2(\Omega)]^2, \sigma \in L^\infty(\Omega) \cap H^1(\Omega). \tag{4.1}$$

We assume the solution has the following regularity properties:

$$y \in H^{r_y}(\Omega), \quad z \in H^{r_z}(\Omega), \quad \boldsymbol{q} \in [H^{r_{\boldsymbol{q}}}(\Omega)]^2, \quad \boldsymbol{p} \in [H^{r_{\boldsymbol{p}}}(\Omega)]^2, \tag{4.2a}$$

$$r_y \geq 1, \quad r_z \geq 2, \quad r_{\boldsymbol{q}} \geq 0, \quad r_{\boldsymbol{p}} \geq 1. \tag{4.2b}$$

Thanks to Theorem 2.1, the condition (4.2) is guaranteed to hold.

It is worthwhile to mention that if $\boldsymbol{q}$ has a well-defined boundary trace in $L^2(\Gamma)$, i.e., $r_{\boldsymbol{q}} > 1/2$, then we refer to this as the high regularity case for the boundary control problem; otherwise, if $r_{\boldsymbol{q}} \in [0, 1/2]$, then we say this is the low regularity case. By Theorem 2.1, if $y_d \in H^{t^*}(\Omega)$ for some $t^* \in (1/2, 1)$, and $\pi/3 \leq \omega < 2\pi/3$, then we are guaranteed to be in the high regularity case. However, if one of the above assumptions concerning y_d or ω is not satisfied, then $\boldsymbol{q}$ is no longer guaranteed to have a well-defined boundary trace.

For the *diffusion dominated* boundary control problem, we gave a rigorous error analysis of a different HDG method for the high regularity case in [33, 32] and for the low regularity case in [25]. In this work, we are interested in the *convection dominated* case. However, existing numerical analysis works for convection dominated diffusion PDEs only consider the high regularity case; see, e.g., [3, 23]. To the best of our knowledge, there is no existing error analysis work on convection dominated PDEs with low regularity solutions.

We now state our main convergence result.

THEOREM 4.1. *Let* $s_y = \min\{r_y, k+1\}$, $s_z = \min\{r_z, k+1\}$, (u, y, z), *and* (u_h, y_h, z_h) *be the solutions of* (1.3) *and* (2.3), *respectively. If assumptions* (A1)–(A3) *hold, then there exists* h_0, *independent of* ε, *such that for all* $\varepsilon < h \leq h_0$ *we have*

$$\|u - u_h\|_{\mathcal{E}_h^\partial} \leq C\left(h^{s_y - 1/2}\|y\|_{s_y} + h^{s_z - 1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),$$

$$\|y - y_h\|_{\mathcal{T}_h} \leq C\left(h^{s_y - 1/2}\|y\|_{s_y} + h^{s_z - 1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),$$

$$\|z - z_h\|_{\mathcal{T}_h} \leq C\left(h^{s_y - 1/2}\|y\|_{s_y} + h^{s_z - 1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),$$

where $\delta(t) = 1$ *if* $t \leq 3/2$, *and otherwise* $\delta(t) = 0$.

Remark 4.2. If $s_y \leq 3/2$, then we have $\|\Delta y\|_{\mathcal{T}_h}$ in the error estimates. This term is finite by Theorem 2.1.

COROLLARY 4.3. *Suppose* $f = 0$, $y_d \in H^{t^*}(\Omega)$ *for some* $t^* \in [0, 1)$, *and assumptions* (A1)–(A3) *hold. Let* $\pi/3 \leq \omega < \pi$ *be the largest interior angle of* Γ, *and let* $r > 0$ *satisfy*

$$r \leq r_d := \frac{1}{2} + t^* \in [1/2, 3/2), \quad \text{and} \quad r < r_\Omega := \min\left\{\frac{3}{2}, \frac{\pi}{\omega} - \frac{1}{2}\right\} \in (1/2, 3/2].$$

If $k=1$, then there exists h_0, independent of ε, such that for all $\varepsilon < h \le h_0$ we have

$$
\begin{aligned}
\|u-u_h\|_{\varepsilon_h^\partial} &\le Ch^r(\|y\|_{H^{r+1/2}(\Omega)} + \|z\|_{H^{r+3/2}(\Omega)} + \delta(r{+}1/2)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}),\\
\|y-y_h\|_{\mathcal{T}_h} &\le Ch^r(\|y\|_{H^{r+1/2}(\Omega)} + \|z\|_{H^{r+3/2}(\Omega)} + \delta(r{+}1/2)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}),\\
\|z-z_h\|_{\mathcal{T}_h} &\le Ch^r(\|y\|_{H^{r+1/2}(\Omega)} + \|z\|_{H^{r+3/2}(\Omega)} + \delta(r{+}1/2)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}).
\end{aligned}
$$

Furthermore, if $k=0$, then there exists h_1, independent of ε, such that for all $\varepsilon < h \le h_1$ we have

$$
\begin{aligned}
\|u-u_h\|_{\varepsilon_h^\partial} &\le Ch^{1/2}(\|y\|_{H^1(\Omega)} + \|z\|_{H^1(\Omega)} + \varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}),\\
\|y-y_h\|_{\mathcal{T}_h} &\le Ch^{1/2}(\|y\|_{H^1(\Omega)} + \|z\|_{H^1(\Omega)} + \varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}),\\
\|z-z_h\|_{\mathcal{T}_h} &\le Ch^{1/2}(\|y\|_{H^1(\Omega)} + \|z\|_{H^1(\Omega)} + \varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}).
\end{aligned}
$$

Similar to [33, 32], the convergence rates are optimal for the control when $k=1$ and suboptimal when $k=0$. However, if $y_d \in L^2(\Omega)$, then $u \in H^{1/2}(\Gamma)$ only and the convergence rate for the control is optimal when $k=0$.

4.2. Proof of Theorem 4.1. We introduce an auxiliary problem with the approximate control u_h in the HDG discretized optimality system (2.6a) replaced by a projection of the exact optimal control and split the proof into seven steps.

We first bound the error between the solution of the optimality system (2.2a)–(2.2d) and $(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u), \boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u)) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$ satisfying the auxiliary problem

$$\mathcal{B}_1(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h(u); \boldsymbol{r}_1, w_1, \widehat{w}_1^o) = -(f, w_1)_{\mathcal{T}_h} - \langle \Pi_k^\partial u, \tau_2 w_1 + \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial}, \tag{4.3a}$$

$$\mathcal{B}_2(\boldsymbol{p}_h(u), z_h(u), \widehat{z}_h(u); \boldsymbol{r}_2, w_2, \widehat{w}_2^o) = -(y_h(u) - y_d, w_2)_{\mathcal{T}_h} \tag{4.3b}$$

for all $(\boldsymbol{r}_1, w_1, \widehat{w}_1^o, \boldsymbol{r}_2, w_2, \widehat{w}_2^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$.

4.2.1. Step 1: Errors between the auxiliary problem (4.3) and the continuous problem (2.2).

LEMMA 4.4. *Let $(\boldsymbol{q}, y, \boldsymbol{p}, z, u)$ be the solution of (2.2). Then for all $(\boldsymbol{r}_1, w_1, \widehat{w}_1^o, \boldsymbol{r}_2, w_2, \widehat{w}_2^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$, we have*

$$
\begin{aligned}
&\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \Pi_k^o y, \widetilde{\Pi}_k^\partial y; \boldsymbol{r}_1, w_1, \widehat{w}_1^o)\\
&\qquad = -(f, w_1)_{\mathcal{T}_h} - \langle \Pi_k^\partial u, \tau_2 w_1 + \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial} + E_1(\boldsymbol{q}, y; w_1, \widehat{w}_1^o),
\end{aligned} \tag{4.4a}
$$

$$\mathcal{B}_2(\boldsymbol{\Pi}_k^o\boldsymbol{p}, \Pi_k^o z, \widetilde{\Pi}_k^\partial z; \boldsymbol{r}_2, w_2, \widehat{w}_2^o) = -(y - y_d, w_2)_{\mathcal{T}_h} + E_2(\boldsymbol{p}, z; w_2, \widehat{w}_2^o), \tag{4.4b}$$

where

$$
\begin{aligned}
E_1(\boldsymbol{q}, y; w_1, \widehat{w}_1^o) = {}& (\Pi_k^o y - y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} - (\sigma(\Pi_k^o - \mathbb{I})y, w_1)_{\mathcal{T}_h}\\
&+ \langle \widehat{w}_1^o - w_1, (\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&+ \langle y - \Pi_k^\partial y, \boldsymbol{\beta}\cdot\boldsymbol{n}(w_1 - \widehat{w}_1^o)\rangle_{\partial\mathcal{T}_h}\\
&- \langle \tau_1(\Pi_k^o y - \Pi_k^\partial y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h},
\end{aligned} \tag{4.5}
$$

$$
\begin{aligned}
E_2(\boldsymbol{p}, z; w_2, \widehat{w}_2^o) = {}& -(\Pi_k^o z - z, \boldsymbol{\beta}\cdot\nabla w_2)_{\mathcal{T}_h} - (\sigma(\Pi_k^o - \mathbb{I})(z + \nabla\cdot\boldsymbol{\beta}), w_2)_{\mathcal{T}_h}\\
&+ \langle \widehat{w}_2^o - w_2, (\boldsymbol{\Pi}_k^o\boldsymbol{p} - \boldsymbol{p})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&- \langle z - \Pi_k^\partial z, \boldsymbol{\beta}\cdot\boldsymbol{n}(w_2 - \widehat{w}_2^o)\rangle_{\partial\mathcal{T}_h}\\
&- \langle \tau_2(\Pi_k^o z - \Pi_k^\partial z), w_2 - \widehat{w}_2^o\rangle_{\partial\mathcal{T}_h}.
\end{aligned} \tag{4.6}
$$

Proof. We only give a proof of (4.4a). By the definition of $\mathcal{B}_1$ in (2.4), one gets

$$
\begin{aligned}
&\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \Pi_k^o y, \widetilde{\Pi}_k^\partial y; \boldsymbol{r}_1, w_1, \widehat{w}_1^o)\\
&\quad= \varepsilon^{-1}(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \boldsymbol{r}_1)_{\mathcal{T}_h} - (\Pi_k^o y, \nabla\cdot\boldsymbol{r}_1)_{\mathcal{T}_h} + \langle\widetilde{\Pi}_k^\partial y, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\qquad - (w_1, \nabla\cdot\boldsymbol{\Pi}_k^o\boldsymbol{q})_{\mathcal{T}_h} + \langle\widehat{w}_1^o, \boldsymbol{\Pi}_k^o\boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\tau_1(\Pi_k^o y - \widetilde{\Pi}_k^\partial y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad + (\Pi_k^o y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} - \langle\widetilde{\Pi}_k^\partial y, \boldsymbol{\beta}\cdot\boldsymbol{n}w_1\rangle_{\partial\mathcal{T}_h} - (\sigma\Pi_k^o y, w_1)_{\mathcal{T}_h}.
\end{aligned}
$$

By the orthogonality properties of $\boldsymbol{\Pi}_k^o$, Π_k^o, $\widetilde{\Pi}_k^\partial$ and the fact $y = u$ on $\mathcal{E}_h^\partial$, we have

$$
\begin{aligned}
&\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \Pi_k^o y, \widetilde{\Pi}_k^\partial y; \boldsymbol{r}_1, w_1, \widehat{w}_1^o)\\
&\quad= \varepsilon^{-1}(\boldsymbol{q}, \boldsymbol{r}_1)_{\mathcal{T}_h} - (y, \nabla\cdot\boldsymbol{r}_1)_{\mathcal{T}_h} + \langle y, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\Pi_k^\partial u, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial}\\
&\qquad + (\nabla w_1, \boldsymbol{q})_{\mathcal{T}_h} + \langle\widehat{w}_1^o - w_1, \boldsymbol{\Pi}_k^o\boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\tau_1(\Pi_k^o y - \widetilde{\Pi}_k^\partial y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad + (\Pi_k^o y - y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} - \langle\Pi_k^\partial y - y, \boldsymbol{\beta}\cdot\boldsymbol{n}w_1\rangle_{\partial\mathcal{T}_h} - (\sigma(\Pi_k^o y - y), w_1)_{\mathcal{T}_h}\\
&\qquad + (y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} - \langle y, \boldsymbol{\beta}\cdot\boldsymbol{n}w_1\rangle_{\partial\mathcal{T}_h} - (\sigma y, w_1)_{\mathcal{T}_h} + \langle\Pi_k^\partial u, \boldsymbol{\beta}\cdot\boldsymbol{n}w_1\rangle_{\mathcal{E}_h^\partial}.
\end{aligned}
$$

By integration by parts, and the fact $\langle\widehat{w}_1^o, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0$, we arrive at

$$
\begin{aligned}
&\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \Pi_k^o y, \widetilde{\Pi}_k^\partial y; \boldsymbol{r}_1, w_1, \widehat{w}_1^o)\\
&\quad= \varepsilon^{-1}(\boldsymbol{q}, \boldsymbol{r}_1)_{\mathcal{T}_h} + (\nabla y, \boldsymbol{r}_1)_{\mathcal{T}_h} - \langle\Pi_k^\partial u, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial} - (w_1, \nabla\cdot\boldsymbol{q})_{\mathcal{T}_h}\\
&\qquad + \langle\widehat{w}_1^o - w_1, (\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\tau_1(\Pi_k^o y - \Pi_k^\partial y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad - \langle\tau_1\Pi_k^\partial u, w_1\rangle_{\mathcal{E}_h^\partial} + (\Pi_k^o y - y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} + \langle y - \Pi_k^\partial y, \boldsymbol{\beta}\cdot\boldsymbol{n}w_1\rangle_{\partial\mathcal{T}_h}\\
&\qquad - (\nabla\cdot(\boldsymbol{\beta}y), w_1)_{\mathcal{T}_h} + \langle\boldsymbol{\beta}\cdot\boldsymbol{n}\Pi_k^\partial u, w_1\rangle_{\mathcal{E}_h^\partial} - (\sigma y, w_1)_{\mathcal{T}_h} - (\sigma(\Pi_k^o - \mathbb{I})y, w_1)_{\mathcal{T}_h}.
\end{aligned}
$$

Then by the facts $\varepsilon^{-1}\boldsymbol{q} = -\nabla y$ and $\nabla\cdot\boldsymbol{q} + \nabla\cdot(\boldsymbol{\beta}y) + \sigma y = f$, we have

$$
\begin{aligned}
&\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}, \Pi_k^o y, \widetilde{\Pi}_k^\partial y; \boldsymbol{r}_1, w_1, \widehat{w}_1^o)\\
&\quad= -(f, w_1)_{\mathcal{T}_h} - \langle\Pi_k^\partial u, \tau_2 w_1 + \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial}\\
&\qquad + (\Pi_k^o y - y, \boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} - (\sigma(\Pi_k^o - \mathbb{I})y, w_1)_{\mathcal{T}_h}\\
&\qquad + \langle\widehat{w}_1^o - w_1, (\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\tau_1(\Pi_k^o y - \Pi_k^\partial y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle y - \Pi_k^\partial y, \boldsymbol{\beta}\cdot\boldsymbol{n}(w_1 - \widehat{w}_1^o)\rangle_{\partial\mathcal{T}_h},
\end{aligned}
\tag{4.7}
$$

where we used $\langle y - \Pi_k^\partial y, \boldsymbol{\beta}\cdot\boldsymbol{n}\widehat{w}_1^o\rangle_{\partial\mathcal{T}_h} = 0$ in (4.7). ∎

By (4.4) and (4.3) we have the following error equations.

LEMMA 4.5. *Let* $(\boldsymbol{q}, y, \boldsymbol{p}, z, u)$ *and* $(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u), \boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u)) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$ *be the solutions of* (2.2) *and* (4.3), *respectively. Then for all* $(\boldsymbol{r}_1, w_1, \widehat{w}_1^o, \boldsymbol{r}_2, w_2, \widehat{w}_2^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$, *we have*

$$
\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q}_h(u), \Pi_k^o y - y_h(u), \widetilde{\Pi}_k^\partial y - \widehat{y}_h^o(u); \boldsymbol{r}_1, w_1, \widehat{w}_1^o) = E_1(\boldsymbol{q}, y; w_1, \widehat{w}_1^o), \tag{4.8a}
$$

$$
\begin{aligned}
\mathcal{B}_2(\boldsymbol{\Pi}_k^o\boldsymbol{p} - \boldsymbol{p}_h(u), \Pi_k^o z - z_h(u), \widetilde{\Pi}_k^\partial z - \widehat{z}_h^o(u); \boldsymbol{r}_2, w_2, \widehat{w}_2^o) &= -(y - y_h(u), w_2)_{\mathcal{T}_h}\\
&\quad + E_2(\boldsymbol{p}, z; w_2, \widehat{w}_2^o),
\end{aligned}
\tag{4.8b}
$$

where E_1 *and* E_2 *are defined in Lemma* 4.4.

LEMMA 4.6. *Let* $(\boldsymbol{q}, y, \boldsymbol{p}, z)$ *be the solution of* (2.2). *Then for all* $(w_1, w_2, \widehat{w}_1^o, \widehat{w}_2^o) \in [W_h \times M_h^o]^2$, *we have*

$$|E_1(\boldsymbol{q}, y; w_1, \widehat{w}_1^o)| \le C\Big(h^{s_y-1/2}\|y\|_{s_y} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\Big)\,\|(w_1, \widehat{w}_1^o)\|_W\,, \tag{4.9}$$

$$|E_2(\boldsymbol{p}, z; w_2, \widehat{w}_2^o)| \le Ch^{s_z-1/2}\|z\|_{s_z}\,\|(w_2, \widehat{w}_2^o)\|_W\,. \tag{4.10}$$

Proof. Since the proof for (4.10) is similar to the proof of (4.9), we only prove (4.9). To simplify notation, we write E_1 from Lemma 4.4 as $E_1(\boldsymbol{q}, y; v_h, \widehat{v}_h^o) = \sum_{i=1}^5 R_i$. For the term R_1, since $((\mathbb{I} - \Pi_k^o)y, \boldsymbol{\Pi}_0^o\boldsymbol{\beta}\cdot\nabla w_1)_{\mathcal{T}_h} = 0$, we get

$$\begin{aligned}
|R_1| &= |((\mathbb{I} - \Pi_k^o)y, (\boldsymbol{\beta} - \boldsymbol{\Pi}_0^o\boldsymbol{\beta})\cdot\nabla w_1)_{\mathcal{T}_h}|\\
&\le Ch^{s_y}|\boldsymbol{\beta}|_{1,\infty}\|y\|_{s_y}\left(\sum_{T\in\mathcal{T}_h} h_T^2\|\nabla w_1\|_T^2\right)^{1/2}\\
&\le Ch^{s_y}\|y\|_{s_y}\|w_1\|_{\mathcal{T}_h} && \text{by (3.5f)}\\
&\le Ch^{s_y}\|y\|_{s_y}\,\|(w_1, \widehat{w}_1^o)\|_W && \text{by (3.14b).}
\end{aligned}$$

For the term R_2, by a direct estimate, we get

$$\begin{aligned}
|R_2| &\le C\|\sigma\|_{0,\infty}^{1/2}\|(\mathbb{I} - \Pi_k^o)y\|_{\mathcal{T}_h}\|\sigma^{1/2}w_1\|_{\mathcal{T}_h}\\
&\le Ch^{s_y}\|\sigma\|_{0,\infty}^{1/2}\|y\|_{s_y}\,\|(w_1, \widehat{w}_1^o)\|_W && \text{by (3.14b)}\\
&\le Ch^{s_y}\|y\|_{s_y}\,\|(w_1, \widehat{w}_1^o)\|_W\,.
\end{aligned}$$

For the term R_3, we need a refined analysis since this term relates to the boundary trace of the gradient of y. Below, we use $\boldsymbol{q} = -\varepsilon\nabla y$ and $\varepsilon < \min_{T\in\mathcal{T}_h}\{h_T\}$.

If $s_y > 3/2$, we have

$$\begin{aligned}
|R_3| &= \varepsilon|\langle \boldsymbol{n}\cdot(\nabla y - \boldsymbol{\Pi}_k^o\nabla y), w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|\\
&\le Ch^{s_y-1}\varepsilon^{1/2}\|y\|_{s_y}\left(\sum_{T\in\mathcal{T}_h}\varepsilon h_T^{-1}\|w_1 - \widehat{w}_1^o\|_{\partial T}^2\right)^{1/2}\\
&\le Ch^{s_y-1/2}\|y\|_{s_y}\,\|(w_1, \widehat{w}_1^o)\|_W\,.
\end{aligned}$$

If $s_y \le 3/2$, use $\langle\boldsymbol{n}\cdot\nabla y, \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h} = 0$ and integration by parts to obtain

$$\begin{aligned}
|R_3| &= \varepsilon|\langle\boldsymbol{n}\cdot\nabla y, w_1\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{n}\cdot\boldsymbol{\Pi}_k^o\nabla y, w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|\\
&= \varepsilon|(\Delta y, w_1)_{\mathcal{T}_h} + (\boldsymbol{\Pi}_k^o\nabla y, \nabla w_1)_{\mathcal{T}_h} - \langle\boldsymbol{n}\cdot\boldsymbol{\Pi}_k^o\nabla y, w_1 - \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|.
\end{aligned}$$

We use integration by parts again and also (3.12b) and (3.12a) to get

$$\begin{aligned}
|R_3| &= \varepsilon|(\Delta y, w_1)_{\mathcal{T}_h} - (\nabla\cdot\boldsymbol{\Pi}_k^o\nabla y, w_1)_{\mathcal{T}_h} + \langle\boldsymbol{n}\cdot\boldsymbol{\Pi}_k^o\nabla y, \widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|\\
&= \varepsilon|(\Delta y, w_1)_{\mathcal{T}_h} - (\nabla\cdot\boldsymbol{\Pi}_k^o\nabla y, \mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h} + \langle\boldsymbol{n}\cdot\boldsymbol{\Pi}_k^o\nabla y, \mathcal{I}_h(w_1, \widehat{w}_1^o)\rangle_{\partial\mathcal{T}_h}|\\
&= \varepsilon|(\Delta y, w_1)_{\mathcal{T}_h} + (\boldsymbol{\Pi}_k^o\nabla y, \nabla\mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h}|.
\end{aligned}$$

Therefore, by the triangle inequality, integration by parts, (3.12c), and (3.12d) we have

$$\begin{aligned}
|R_3| &\le \varepsilon|(\Delta y, \mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h} + (\boldsymbol{\Pi}_k^o\nabla y, \mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h}| + \varepsilon|(\Delta y, w_1 - \mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h}|\\
&= \varepsilon|(\nabla y - \boldsymbol{\Pi}_k^o\nabla y, \nabla\mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h}| + \varepsilon|(\Delta y, w_1 - \mathcal{I}_h(w_1, \widehat{w}_1^o))_{\mathcal{T}_h}|\\
&\le C\varepsilon^{1/2}(h^{s_y-1}\|y\|_{s_y} + h\|\Delta y\|_{\mathcal{T}_h})\,\|(w_1, \widehat{w}_1^o)\|_W\\
&\le C(h^{s_y-1/2}\|y\|_{s_y} + \varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h})\,\|(w_1, \widehat{w}_1^o)\|_W\,.
\end{aligned}$$

For the terms R_4 and R_5, use the Cauchy–Schwarz inequality and (3.2) to get

$$
\begin{aligned}
|R_4| &= |\langle \boldsymbol{\beta}\cdot\boldsymbol{n}(y-\Pi_k^\partial y), w_1-\widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|\\
&\le Ch^{s_y-1/2}\|\boldsymbol{\beta}\|_{0,\infty}^{1/2}\|y\|_{s_y}\left(\sum_{T\in\mathcal{T}_h}|\boldsymbol{\beta}\cdot\boldsymbol{n}|^{1/2}(w_1-\widehat{w}_1^o)\|_{\partial T}^2\right)^{1/2}\\
&\le Ch^{s_y-1/2}\|y\|_{s_y}\,\|(w_1,\widehat{w}_1^o)\|_W\,,\\
|R_5| &= |\langle \tau_1(\Pi_k^o-\Pi_k^\partial)y, w_1-\widehat{w}_1^o\rangle_{\partial\mathcal{T}_h}|\\
&\le Ch^{s_y-1}(\beta_0^{1/2}h^{1/2}+\varepsilon^{1/2})\|y\|_{s_y}\,\|(w_1,\widehat{w}_1^o)\|_W\\
&\le Ch^{s_y-1/2}\|y\|_{s_y}\,\|(w_1,\widehat{w}_1^o)\|_W\,.
\end{aligned}
$$

From all the estimates above we get our final result. ∎

LEMMA 4.7. *Let* $(\boldsymbol{q},y,\boldsymbol{p},z)$ *and* $(\boldsymbol{q}_h(u),y_h(u),\widehat{y}_h^o(u),\,\boldsymbol{p}_h(u),z_h(u),\widehat{z}_h^o(u))\in[\boldsymbol{V}_h\times W_h\times M_h^o]^2$ *be the solutions of* (2.2) *and* (4.3), *respectively. If assumptions* (A1) *and* (A2) *hold, then there exists* $h_0>0$, *independent of* ε, *such that for all* $\varepsilon<h\le h_0$ *we have the error estimates*

$$
\begin{aligned}
\|y-y_h(u)\|_{\mathcal{T}_h} &\le C\left(h^{s_y-1/2}\|y\|_{s_y}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),\\
\|z-z_h(u)\|_{\mathcal{T}_h} &\le C\left(h^{s_y-1/2}\|y\|_{s_y}+h^{s_z-1/2}\|z\|_{s_z}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),\\
\|\boldsymbol{p}-\boldsymbol{p}_h(u)\|_{\mathcal{T}_h} &\le C\varepsilon^{1/2}\left(h^{s_y-1/2}\|y\|_{s_y}+h^{s_z-1/2}\|z\|_{s_z}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}
$$

Proof. By Theorem 3.14, (4.8a), and (4.9) we get

$$
\begin{aligned}
&\|(\boldsymbol{\Pi}_k^o\boldsymbol{q}-\boldsymbol{q}_h(u),\Pi_k^o y-y_h(u),\widetilde{\Pi}_k^\partial y-\widehat{y}_h^o(u))\|\\
&\quad\le C\sup_{0\neq(\boldsymbol{r}_1,w_1,\widehat{w}_1^o)\in\boldsymbol{V}_h\times W_h\times M_h^o}\frac{\mathcal{B}_1(\boldsymbol{\Pi}_k^o\boldsymbol{q}-\boldsymbol{q}_h(u),\Pi_k^o y-y_h(u),\widetilde{\Pi}_k^\partial y-\widehat{y}_h^o(u);\boldsymbol{r}_1,w_1,\widehat{w}_1^o)}{\|(\boldsymbol{r}_1,w_1,\widehat{w}_1^o)\|}\\
&\quad\le C\sup_{0\neq(\boldsymbol{r}_1,w_1,\widehat{w}_1^o)\in\boldsymbol{V}_h\times W_h\times M_h^o}\frac{E_1(\boldsymbol{q},y;v_h,\widehat{v}_h^o)}{\|(\boldsymbol{r}_1,w_1,\widehat{w}_1^o)\|}\\
&\quad\le C\left(h^{s_y-1/2}\|y\|_{s_y}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&|y-y_h(u)\|_{\mathcal{T}_h}\\
&\quad\le\|y-\Pi_k^o y\|_{\mathcal{T}_h}+\|\Pi_k^o y-y_h(u)\|_{\mathcal{T}_h}\\
&\quad\le\|y-\Pi_k^o y\|_{\mathcal{T}_h}+C\|(\boldsymbol{\Pi}_k^o\boldsymbol{q}-\boldsymbol{q}_h(u),\Pi_k^o y-y_h(u),\widetilde{\Pi}_k^\partial y-\widehat{y}_h^o(u))\|\\
&\quad\le C\left(h^{s_y-1/2}\|y\|_{s_y}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}\tag{4.11}
$$

By Theorem 3.14, (4.8b), (4.10), and estimate (4.11) we get

$$
\begin{aligned}
&\|(\boldsymbol{\Pi}_k^o\boldsymbol{p}-\boldsymbol{p}_h(u),\Pi_k^o z-z_h(u),\widetilde{\Pi}_k^\partial z-\widehat{z}_h^o(u))\|\\
&\quad\le C\sup_{0\neq(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\in\boldsymbol{V}_h\times W_h\times M_h^o}\frac{\mathcal{B}_2(\boldsymbol{\Pi}_k^o\boldsymbol{p}-\boldsymbol{p}_h(u),\Pi_k^o z-z_h(u),\widetilde{\Pi}_k^\partial z-\widehat{z}_h^o(u);\boldsymbol{r}_2,w_2,\widehat{w}_2^o)}{\|(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\|}\\
&\quad\le C\sup_{0\neq(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\in\boldsymbol{V}_h\times W_h\times M_h^o}\frac{E_2(\boldsymbol{p},z;w_2,\widehat{w}_2^o)-(y-y_h(u),w_2)}{\|(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\|}\\
&\quad\le C\left(h^{s_z-1/2}\|z\|_{s_z}+h^{s_y-1/2}\|y\|_{s_y}+\delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}
$$

Therefore, the triangle inequality gives

$$\begin{aligned}
&\|z - z_h(u)\|_{\mathcal{T}_h}\\
&\quad \le \|z - \Pi_k^o z\|_{\mathcal{T}_h} + \|\Pi_k^o z - z_h(u)\|_{\mathcal{T}_h}\\
&\quad \le \|z - \Pi_k^o z\|_{\mathcal{T}_h} + (\beta_0 + \sigma_0)^{-1} \|(\boldsymbol{\Pi}_k^o \boldsymbol{p} - \boldsymbol{p}_h(u), \Pi_k^o z - z_h(u), \widetilde{\Pi}_k^\partial z - \widehat{z}_h^o(u))\|\\
&\quad \le C\left(h^{s_z - 1/2}\|z\|_{s_z} + h^{s_y - 1/2}\|y\|_{s_y} + \delta(s_y)\varepsilon^{1/2} h \|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}$$

Next, we use the triangle inequality, $\boldsymbol{p} = -\varepsilon \nabla z$, and $\varepsilon < \min_{T\in\mathcal{T}_h}\{h_T\}$ to get

$$\begin{aligned}
\|\boldsymbol{p} - \boldsymbol{p}_h(u)\|_{\mathcal{T}_h} &\le \|\boldsymbol{p} - \boldsymbol{\Pi}_k^o \boldsymbol{p}\|_{\mathcal{T}_h} + \|\boldsymbol{\Pi}_k^o \boldsymbol{p} - \boldsymbol{p}_h(u)\|_{\mathcal{T}_h}\\
&\le \|\boldsymbol{p} - \boldsymbol{\Pi}_k^o \boldsymbol{p}\|_{\mathcal{T}_h} + \varepsilon^{1/2}\|(\boldsymbol{\Pi}_k^o \boldsymbol{p} - \boldsymbol{p}_h(u), \Pi_k^o z - z_h(u), \widetilde{\Pi}_k^\partial z - \widehat{z}_h^o(u))\|\\
&\le C\varepsilon^{1/2}\left(h^{s_z - 1/2}\|z\|_{s_z} + h^{s_y - 1/2}\|y\|_{s_y} + \delta(s_y)\varepsilon^{1/2} h\|\Delta y\|_{\mathcal{T}_h}\right). \qquad \square
\end{aligned}$$

4.2.2. Step 2: Errors between the auxiliary problem (4.3) and the discrete problem (2.6). By (2.6) and (4.3) we have the following error equations.

LEMMA 4.8. *Let* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{p}_h, z_h, \widehat{z}_h^o, u_h) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2 \times M_h^\partial$ *and* $(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u), \boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u)) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$ *be the solutions of* (2.6) *and* (4.3), *respectively. Then for all* $(\boldsymbol{r}_1, w_1, \widehat{w}_1^o, \boldsymbol{r}_2, w_2, \widehat{w}_2^o) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$, *we have*

$$\begin{aligned}
&\mathcal{B}_1(\boldsymbol{q}_h - \boldsymbol{q}_h(u), y_h - y_h(u), \widehat{y}_h^o - \widehat{y}_h^o(u); \boldsymbol{r}_1, w_1, \widehat{w}_1^o) =\\
&\qquad\qquad -\langle u_h - \Pi_k^\partial u, \tau_2 w_1 + \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\mathcal{E}_h^\partial}, \qquad (4.12a)
\end{aligned}$$

$$\mathcal{B}_2(\boldsymbol{p}_h - \boldsymbol{p}_h(u), z_h - z_h(u), \widehat{z}_h^o - \widehat{z}_h^o(u); \boldsymbol{r}_2, w_2, \widehat{w}_2^o) = -(y_h - y_h(u), w_2)_{\mathcal{T}_h}. \qquad (4.12b)$$

LEMMA 4.9. *Let* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{p}_h, z_h, \widehat{z}_h^o, u_h) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2 \times M_h^\partial$ *and*

$$(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u), \boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u)) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$$

be the solutions of (2.6) *and* (4.3), *respectively. Then we have*

$$\begin{aligned}
&\|y_h - y_h(u)\|_{\mathcal{T}_h}^2 + \gamma\|u_h - \Pi_k^\partial u\|_{\mathcal{E}_h^\partial}^2\\
&\qquad = \langle u_h - \Pi_k^\partial u, \Pi_k^\partial(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{p}_h(u)\cdot\boldsymbol{n} - \tau_2(z_h(u) - \widehat{z}_h^o(u))\rangle_{\mathcal{E}_h^\partial}.
\end{aligned}$$

Proof. Take $(\boldsymbol{r}_1, w_1, \widehat{w}_1^o) = (\boldsymbol{p}_h - \boldsymbol{p}_h(u), z_h - z_h(u), \widehat{z}_h^o - \widehat{z}_h^o(u))$ and $(\boldsymbol{r}_2, w_2, \widehat{w}_2^o) = (\boldsymbol{q}_h - \boldsymbol{q}_h(u), y_h - y_h(u), \widehat{y}_h^o - \widehat{y}_h^o(u))$ in (4.12a) and (4.12b), respectively, and use Lemma 2.3 to get

$$\begin{aligned}
&-\|y_h - y_h(u)\|_{\mathcal{T}_h}^2\\
&\quad = \mathcal{B}_2(\boldsymbol{p}_h - \boldsymbol{p}_h(u), z_h - z_h(u), \widehat{z}_h^o - \widehat{z}_h^o(u); \boldsymbol{q}_h - \boldsymbol{q}_h(u), y_h - y_h(u), \widehat{y}_h^o - \widehat{y}_h^o(u))\\
&\quad = \mathcal{B}_1(\boldsymbol{q}_h - \boldsymbol{q}_h(u), y_h - y_h(u), \widehat{y}_h^o - \widehat{y}_h^o(u); \boldsymbol{p}_h - \boldsymbol{p}_h(u), z_h - z_h(u), \widehat{z}_h^o - \widehat{z}_h^o(u))\\
&\quad = -\langle u_h - \Pi_k^\partial u, (\boldsymbol{p}_h - \boldsymbol{p}_h(u))\cdot\boldsymbol{n} + \tau_2(z_h - z_h(u))\rangle_{\mathcal{E}_h^\partial}.
\end{aligned}$$

Therefore, (2.3e), (2.2), and $\widehat{z}_h^o = \widehat{z}_h^o(u) = 0$ on $\mathcal{E}_h^\partial$ give

$$\begin{aligned}
&\|y_h - y_h(u)\|_{\mathcal{T}_h}^2 + \gamma\|u_h - \Pi_k^\partial u\|_{\mathcal{E}_h^\partial}^2\\
&\quad = \langle u_h - \Pi_k^\partial u, \boldsymbol{p}_h\cdot\boldsymbol{n} + \tau_2 z_h + \gamma u_h\rangle_{\mathcal{E}_h^\partial}\\
&\qquad - \langle u_h - \Pi_k^\partial u, \boldsymbol{p}_h(u)\cdot\boldsymbol{n} + \tau_2 z_h(u) + \gamma\Pi_k^\partial u\rangle_{\mathcal{E}_h^\partial}\\
&\quad = \langle u_h - \Pi_k^\partial u, \Pi_k^\partial(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{p}_h(u)\cdot\boldsymbol{n} - \tau_2(z_h(u) - \widehat{z}_h^o(u))\rangle_{\mathcal{E}_h^\partial}. \qquad \square
\end{aligned}$$

LEMMA 4.10. *Let* $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o, \boldsymbol{p}_h, z_h, \widehat{z}_h^o, u_h) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2 \times M_h^\partial$ *and* $(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u), \boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u)) \in [\boldsymbol{V}_h \times W_h \times M_h^o]^2$ *be the solutions of* (2.6) *and* (4.3), *respectively. If assumptions* (A1)–(A3) *hold, then there exists* h_0, *independent of* ε, *such that for all* $\varepsilon < h \leq h_0$, *we have the estimates*

$$\begin{aligned}
\|u - u_h\|_{\mathcal{E}_h^\partial} &\leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),\\
\|y_h - y_h(u)\|_{\mathcal{T}_h} &\leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right),\\
\|z_h - z_h(u)\|_{\mathcal{T}_h} &\leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}$$

Proof. By Lemma 4.9, the Cauchy–Schwarz inequality, and Young's inequality one gets

$$\begin{aligned}
&\|y_h - y_h(u)\|_{\mathcal{T}_h} + \gamma^{1/2}\|u_h - \Pi_k^\partial u\|_{\mathcal{E}_h^\partial}\\
&\qquad \leq C\|\Pi_k^\partial(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{p}_h(u)\cdot\boldsymbol{n}\|_{\mathcal{E}_h^\partial} + C\|\tau_2(z_h(u) - \widehat{z}_h^o(u)\|_{\mathcal{E}_h^\partial}.
\end{aligned}$$

By the triangle inequality, $\boldsymbol{p} = -\varepsilon\nabla z$, an inverse inequality, and the estimates in Lemma 4.7 we have

$$\begin{aligned}
&\|\boldsymbol{\Pi}_k^\partial(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{p}_h(u)\cdot\boldsymbol{n}\|_{\mathcal{E}_h^\partial}\\
&\quad \leq \|\boldsymbol{\Pi}_k^\partial(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{\Pi}_k^o(\boldsymbol{p}\cdot\boldsymbol{n})\|_{\mathcal{E}_h^\partial} + \|\boldsymbol{\Pi}_k^o(\boldsymbol{p}\cdot\boldsymbol{n}) - \boldsymbol{p}_h(u)\cdot\boldsymbol{n}\|_{\mathcal{E}_h^\partial}\\
&\qquad \leq C\varepsilon h^{s_z-3/2}\|z\|_{s_z} + \sum_{T\in\mathcal{T}_h, \partial T\bigcap\mathcal{E}_h^\partial\neq\emptyset} h_T^{-1/2}\|\boldsymbol{\Pi}_k^o\boldsymbol{p} - \boldsymbol{p}_h(u)\|_T\\
&\qquad \leq C\varepsilon h^{s_z-3/2}\|z\|_{s_z} + \sum_{T\in\mathcal{T}_h, \partial T\bigcap\mathcal{E}_h^\partial\neq\emptyset} h_T^{-1/2}\left(\|\boldsymbol{\Pi}_k^o\boldsymbol{p} - \boldsymbol{p}\|_T + \|\boldsymbol{p} - \boldsymbol{p}_h(u)\|_T\right)\\
&\qquad \leq Ch^{s_z-1/2}\|z\|_{s_z} + \varepsilon^{-1/2}\sum_{T\in\mathcal{T}_h, \partial T\bigcap\mathcal{E}_h^\partial\neq\emptyset} \|\boldsymbol{p} - \boldsymbol{p}_h(u)\|_T\\
&\qquad \leq Ch^{s_z-1/2}\|z\|_{s_z} + \varepsilon^{-1/2}\|\boldsymbol{p} - \boldsymbol{p}_h(u)\|_{\mathcal{T}_h}\\
&\qquad \leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}$$

By the triangle inequality, $z = 0$ on Γ, and the estimate (10) we have

$$\begin{aligned}
&\|\tau_2(z_h(u) - \widehat{z}_h^o(u)\|_{\mathcal{E}_h^\partial}\\
&\quad \leq C\|(z_h(u) - \widehat{z}_h^o(u) - \Pi_k^o z + \widetilde{\Pi}_k^\partial z)\|_{\mathcal{E}_h^\partial} + C\|(\Pi_k^o z - \widetilde{\Pi}_k^\partial z)\|_{\mathcal{E}_h^\partial}\\
&\quad \leq C\|(\boldsymbol{\Pi}_k^o\boldsymbol{p} - \boldsymbol{p}_h(u), \Pi_k^o z - z_h(u), \widetilde{\Pi}_k^\partial z - \widehat{z}_h^o(u))\| + Ch^{s_z-1/2}\|z\|_{s_z}\\
&\quad \leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}$$

This implies

$$\begin{aligned}
&\|y_h - y_h(u)\|_{\mathcal{T}_h} + \|u_h - \Pi_k^\partial u\|_{\mathcal{E}_h^\partial}\\
&\quad \leq C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).
\end{aligned}$$

TABLE 1
Example 1: *Smooth test with $k = 1$ and $\varepsilon = 10^{-7}$; errors for the control u, state y, and the adjoint state z.*

$h/\sqrt{2}$	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$\|y - y_h\|_{0,\Omega}$	6.0299E-02	1.3188E-02	2.1788E-03	4.8975E-04	1.1863E-04
order	-	2.1929	2.5976	2.1534	2.0456
$\|z - z_h\|_{0,\Omega}$	1.0572E-01	2.6724E-02	6.2451E-03	1.5091E-03	3.7092E-04
order	-	1.9841	2.0973	2.0491	2.0245
$\|u - u_h\|_{0,\Gamma}$	2.5537E-01	5.6029E-02	1.2108E-02	2.8176E-03	6.7424E-04
order	-	2.1883	2.2102	2.1034	2.0631

By the triangle inequality and the fact $y = u$ on $\mathcal{E}_h^\partial$, we get

$$\begin{aligned}\|u - u_h\|_{\mathcal{E}_h^\partial} &\le \|y - \Pi_k^\partial y\|_{\mathcal{E}_h^\partial} + \|\Pi_k^\partial u - u_h\|_{\mathcal{E}_h^\partial}\\ &\le C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).\end{aligned}$$

By Theorem 3.14 and (4.12b), one has

$$\begin{aligned}&\|(\boldsymbol{p}_h(u) - \boldsymbol{p}_h, z_h(u) - z_h, \widehat{z}_h^o(u) - \widehat{z}_h^o)\|\\ &\quad\le C \sup_{0\neq(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\in \boldsymbol{V}_h\times W_h\times M_h^o} \frac{\mathcal{B}_2(\boldsymbol{p}_h(u) - \boldsymbol{p}_h, z_h(u) - z_h, \widehat{z}_h^o(u) - \widehat{z}_h^o; \boldsymbol{r}_2, w_2, \widehat{w}_2^o)}{\|(\boldsymbol{r}_2, w_2, \widehat{w}_2^o)\|}\\ &\quad\le C \sup_{0\neq(\boldsymbol{r}_2,w_2,\widehat{w}_2^o)\in \boldsymbol{V}_h\times W_h\times M_h^o} \frac{(y_h - y_h(u), w_2)_{\mathcal{T}_h}}{\|(\boldsymbol{r}_2, w_2, \widehat{w}_2^o)\|}\\ &\quad\le C\|y_h - y_h(u)\|_{\mathcal{T}_h}\\ &\quad\le C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right).\end{aligned}$$

Therefore,

$$\|z_h(u) - z_h\|_{\mathcal{T}_h} \le C\left(h^{s_y-1/2}\|y\|_{s_y} + h^{s_z-1/2}\|z\|_{s_z} + \delta(s_y)\varepsilon^{1/2}h\|\Delta y\|_{\mathcal{T}_h}\right). \qquad \square$$

5. Numerical experiments. In this section, we report numerical experiments to illustrate our theoretical results. For all experiments, we take $\gamma = 1$, and the stabilization functions are chosen as in (2.3f)–(2.3g).

5.1. Example 1: Smooth test. In our first test, we take $\Omega = [0,1]\times[0,1] \subset \mathbb{R}^2$, the state, dual state, and convection coefficient are chosen as

$$\begin{gathered} y = -\varepsilon\pi(\sin(\pi x_1) + \sin(\pi x_2)),\ z = \sin(\pi x_1)\sin(\pi x_2),\\ \boldsymbol{\beta} = -[x_1^2\sin(x_2), \cos(x_1)e^{x_2}],\ \sigma = 0,\end{gathered}$$

and the source term f and the desired state y_d are generated using the optimality system (1.3) with the above data. We show the numerical results for $k = 1$ and $\varepsilon = 10^{-7}$ in Table 1.

TABLE 2

Example 2: Nonsmooth test on a square domian with $k = 1$ and different ε; errors for the control u, state y, and the adjoint state z.

ε	$\frac{h}{\sqrt{2}}$	$\Vert y - y_h\Vert_{\mathcal{T}_h}$		$\Vert z - z_h\Vert_{\mathcal{T}_h}$		$\Vert u - u_h\Vert_{\varepsilon_h^\partial}$	
		Error	Rate	Error	Rate	Error	Rate
1/10	2^{-1}	4.8530E-04		1.0783E-03		2.0808E-03	
	2^{-2}	1.6439E-04	1.56	5.3340E-04	1.01	7.2793E-04	1.51
	2^{-3}	5.5062E-05	1.58	2.0652E-04	1.37	2.6827E-04	1.44
	2^{-4}	1.5885E-05	1.79	6.4955E-05	1.67	9.1834E-05	1.54
	2^{-5}	4.1123E-06	1.95	1.7687E-05	1.88	2.7451E-05	1.74
1/50	2^{-1}	1.3406E-03		2.2114E-03		4.1984E-03	
	2^{-2}	4.9404E-04	1.44	1.3080E-03	0.75	1.8165E-03	1.21
	2^{-3}	2.4238E-04	1.02	7.3661E-04	0.83	8.9345E-04	1.02
	2^{-4}	1.1777E-04	1.04	3.5878E-04	1.04	5.2980E-04	0.75
	2^{-5}	4.1387E-05	1.50	1.4389E-04	1.32	2.6837E-04	0.98
1/100	2^{-1}	1.6216E-03		2.5742E-03		4.7465E-03	
	2^{-2}	6.6643E-04	1.28	1.5857E-03	0.70	2.1124E-03	1.17
	2^{-3}	2.8129E-04	1.24	9.2685E-04	0.77	1.0435E-03	1.02
	2^{-4}	1.7254E-04	0.71	5.2512E-04	0.82	7.3691E-04	0.50
	2^{-5}	8.5868E-05	1.00	2.5477E-04	1.04	5.0212E-04	0.53
1/1000	2^{-1}	1.8500E-03		2.9440E-03		5.3059E-03	
	2^{-2}	8.3498E-04	1.14	2.0237E-03	0.54	2.4554E-03	1.11
	2^{-3}	3.9689E-04	1.07	1.3617E-03	0.57	1.0865E-03	1.17
	2^{-4}	2.7140E-04	0.55	9.1122E-04	0.58	5.2906E-04	1.03
	2^{-5}	1.5180E-04	0.84	5.8115E-04	0.65	4.5760E-04	0.21
1/10000	2^{-1}	1.9431E-03		3.0087E-03		5.3675E-03	
	2^{-2}	8.4727E-04	1.20	2.0778E-03	0.53	2.5082E-03	1.10
	2^{-3}	3.9498E-04	1.10	1.4261E-03	0.54	1.1393E-03	1.14
	2^{-4}	2.3736E-04	0.73	9.9958E-04	0.51	5.6195E-04	1.01
	2^{-5}	1.5931E-04	0.58	6.9841E-04	0.52	3.2155E-04	0.81

5.2. Example 2: Nonsmooth test on a square. Next, we take $\Omega = [0,1] \times [0,1] \subset \mathbb{R}^2$ and choose the data as

$$y_d = x(1-x)y(1-y), \quad f = 0, \quad \boldsymbol{\beta} = -[x_1^2 \sin(x_2), \cos(x_1)e^{x_2}], \quad \text{and} \quad \sigma = 0.$$

We tested 5 cases with $k = 1$ and different values for ε and we do not have exact solutions for these problems; we solved the problems numerically for a triangulation with approximately 1.5 million elements and compared these reference solutions against other solutions computed on meshes with larger h. The numerical results are shown in Table 2; the computed convergence rates are erratic and do not follow a clear pattern. The same phenomenon has been observed in another work on a convection dominated Dirichlet boundary control problem [5]. Also, we plot the state and boundary control in Figure 1. Furthermore, many works on convection dominated PDEs observe well-behaved convergence if they remove a small portion of the domain containing the layer; see [31, section 6] for a convection dominated distributed optimal control problem and [23, Table 4, section 5.4] for a convection dominated PDE. We did not compute the rates by removing the layer since the layer is always on the boundary; see Figure 1.

5.3. Example 3: Nonsmooth test on nonsquare domains. Next, we choose the data as

$$\varepsilon = 1/10000, \quad y_d = 1, \quad f = 0, \quad \boldsymbol{\beta} = [1,1], \quad \text{and} \quad \sigma = 0.$$

We tested two cases with different domains and we did not compute the convergence rates. These problems are solved on a triangulation with $h = \sqrt{2}/2^{10}$. We plot the

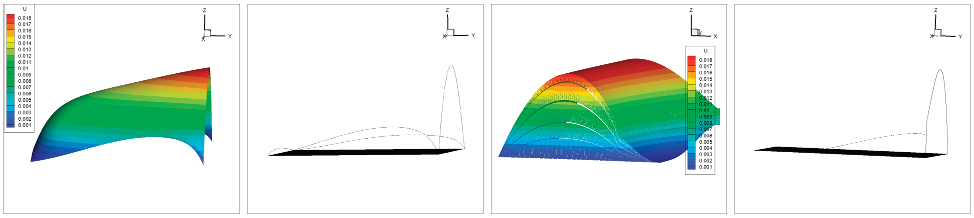

FIG. 1. *Example* 2: *The first two are the state y_h and the control u_h for $\varepsilon = 1/10$; the last two are the state y_h and the control u_h for $\varepsilon = 1/10000$.*

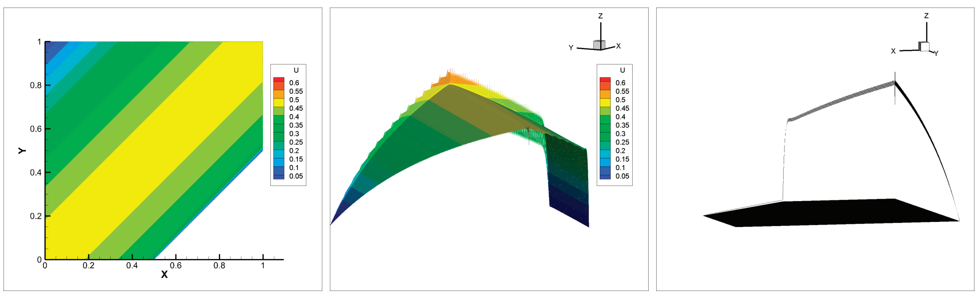

FIG. 2. *Example* 3: *Left and middle are the two-dimensional and three-dimensional plot of the state y_h, respectively; right is the control u_h.*

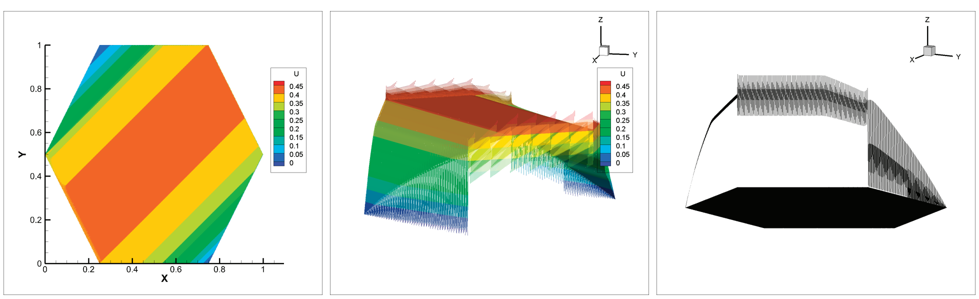

FIG. 3. *Example* 3: *Left and middle are the two-dimensional and three-dimensional plot of the state y_h, respectively; right is the control u_h.*

computed state and boundary control in Figures 2 and 3. We note that the computed control has oscillations, in contrast to the problem considered on the square domain. These tests show that the HDG method has promise for the convection dominated Dirichlet boundary control problem.

6. Conclusion. In [32, 25], we studied an HDG method for a *diffuison dominated* convection diffusion Dirichlet boundary control problem. We obtained optimal convergence rates for the control under a high regularity assumption in [32] and a low regularity assumption in [25]. In this work, we considered a different HDG method with a lower computational cost for a convection dominated convection diffusion boundary control problem under high and low regularity conditions and again proved optimal convergence rates for the control. All existing numerical analysis work on Dirichlet boundary control problems has assumed the mesh is quasi-uniform; however, we do not need to have this assumption here.

To the best of our knowledge, this work is the only existing numerical analysis exploration of this convection dominated diffusion Dirichlet control problem.

REFERENCES

[1] T. Apel, M. Mateos, J. Pfefferer, and A. Rösch, *On the regularity of the solutions of Dirichlet optimal control problems in polygonal domains*, SIAM J. Control Optim., 53 (2015), pp. 3620–3641, https://doi.org/10.1137/140994186.

[2] T. Apel, M. Mateos, J. Pfefferer, and A. Rösch, *Error estimates for Dirichlet control problems in polygonal domains: Quasi-uniform meshes*, Math. Control Relat. Fields, 8 (2018).

[3] B. Ayuso and L. D. Marini, *Discontinuous Galerkin methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1391–1420, https://doi.org/10.1137/080719583.

[4] R. Becker and B. Vexler, *Optimal control of the convection-diffusion equation using stabilized finite element methods*, Numer. Math., 106 (2007), pp. 349–367, https://doi.org/10.1007/s00211-007-0067-0.

[5] P. Benner and H. Yücel, *A Local Discontinuous Galerkin Method for Dirichlet Boundary Control Problems*, https://www.researchgate.net/publication/316862036.

[6] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, 3rd ed., Texts Appl. Math., Springer, New York, 2008, https://doi.org/10.1007/978-0-387-75934-0.

[7] A. Buffa, T. J. R. Hughes, and G. Sangalli, *Analysis of a multiscale discontinuous Galerkin method for convection-diffusion problems*, SIAM J. Numer. Anal., 44 (2006), pp. 1420–1440, https://doi.org/10.1137/050640382.

[8] E. Burman and P. Hansbo, *Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 1437–1453, https://doi.org/10.1016/j.cma.2003.12.032.

[9] E. Casas and J. P. Raymond, *Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations*, SIAM J. Control Optim., 45 (2006), pp. 1586–1611, https://doi.org/10.1137/050626600.

[10] P. Castillo, B. Cockburn, D. Schötzau, and C. Schwab, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp., 71 (2002), pp. 455–478, https://doi.org/10.1090/S0025-5718-01-01317-5.

[11] J. Česenek and M. Feistauer, *Theory of the space-time discontinuous Galerkin method for nonstationary parabolic problems with nonlinear convection and diffusion*, SIAM J. Numer. Anal., 50 (2012), pp. 1181–1206, https://doi.org/10.1137/110828903.

[12] A. Cesmelioglu, B. Cockburn, and W. Qiu, *Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier-Stokes equations*, Math. Comp., 86 (2017), pp. 1643–1670, https://doi.org/10.1090/mcom/3195.

[13] Y. Chen, B. Cockburn, and B. Dong, *Superconvergent HDG methods for linear, stationary, third-order equations in one-space dimension*, Math. Comp., 85 (2016), pp. 2715–2742, https://doi.org/10.1090/mcom/3091.

[14] Y. Cheng, X. Meng, and Q. Zhang, *Application of generalized Gauss-Radau projections for the local discontinuous Galerkin method for linear convection-diffusion equations*, Math. Comp., 86 (2017), pp. 1233–1267, https://doi.org/10.1090/mcom/3141.

[15] Y. Cheng and C. W. Shu, *Superconvergence of discontinuous Galerkin and local discontinuous Galerkin schemes for linear hyperbolic and convection-diffusion equations in one space dimension*, SIAM J. Numer. Anal., 47 (2010), pp. 4044–4072, https://doi.org/10.1137/090747701.

[16] 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 (2009), pp. 1319–1365, https://doi.org/10.1137/070706616.

[17] B. Cockburn, J. Gopalakrishnan, N. Nguyen, J. Peraire, and F. Sayas, *Analysis of HDG methods for Stokes flow*, Math. Comp., 80 (2011), pp. 723–760, https://doi.org/10.1090/S0025-5718-2010-02410-X.

[18] B. Cockburn and K. Mustapha, *A hybridizable discontinuous Galerkin method for fractional diffusion problems*, Numer. Math., 130 (2015), pp. 293–314, https://doi.org/10.1007/s00211-014-0661-x.

[19] B. Cockburn and J. Shen, *A hybridizable discontinuous Galerkin method for the p-Laplacian*, SIAM J. Sci. Comput., 38 (2016), pp. A545–A566, https://doi.org/10.1137/15M1008014.

[20] B. Cockburn and C. W. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463, https://doi.org/10.1137/S0036142997316712.

[21] W. Eckhaus, *Boundary layers in linear elliptic singular perturbation problems*, SIAM Rev., 14 (1972), pp. 225–270, https://doi.org/10.1137/1014030.

[22] A. Ern and J.-L. Guermond, *Finite element quasi-interpolation and best approximation*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 1367–1385.

[23] G. Fu, W. Qiu, and W. Zhang, *An analysis of HDG methods for convection-dominated diffusion problems*, ESAIM Math. Model. Numer. Anal., 49 (2015), pp. 225–256, https://doi.org/10.1051/m2an/2014032.

[24] A. V. Fursikov, M. D. Gunzburger, and L. Hou, *Boundary value problems and optimal boundary control for the Navier–Stokes system: The two-dimensional case*, SIAM J. Control Optim., 36 (1998), pp. 852–894, https://doi.org/10.1137/S0363012994273374.

[25] W. Gong, W. Hu, M. Mateos, J. Singler, X. Zhang, and Y. Zhang, *A New HDG Method for Dirichlet Boundary Control of Convection Diffusion PDEs* II: *Low Regularity*, SIAM J. Numer. Anal., 56 (2018), pp. 2262–2287, https://doi.org/10.1137/17M1152103.

[26] W. Gong and N. Yan, *Mixed finite element method for Dirichlet boundary control problem governed by elliptic PDEs*, SIAM J. Control Optim., 49 (2011), pp. 984–1014, https://doi.org/10.1137/100795632.

[27] M. D. Gunzburger, L. Hou, and T. P. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with Dirichlet controls*, RAIRO Modél. Math. Anal. Numér., 25 (1991), pp. 711–748, https://doi.org/10.1051/m2an/1991250607111.

[28] M. D. Gunzburger, L. Hou, and T. P. Svobodny, *Boundary velocity control of incompressible flow with an application to viscous drag reduction*, SIAM J. Control Optim., 30 (1992), pp. 167–181, https://doi.org/10.1137/0330011.

[29] M. D. Gunzburger and S. Manservisi, *The velocity tracking problem for Navier-Stokes flows with bounded distributed controls*, SIAM J. Control Optim., 37 (1999), pp. 1913–1945, https://doi.org/10.1137/S0363012998337400.

[30] M. D. Gunzburger and S. Manservisi, *Analysis and approximation of the velocity tracking problem for Navier–Stokes flows with distributed control*, SIAM J. Numer. Anal., 37 (2000), pp. 1481–1512, https://doi.org/10.1137/S0036142997329414.

[31] M. Heinkenschloss and D. Leykekhman, *Local error estimates for SUPG solutions of advection-dominated elliptic linear-quadratic optimal control problems*, SIAM J. Numer. Anal., 47 (2010), pp. 4607–4638, https://doi.org/10.1137/090759902.

[32] W. Hu, M. Mateos, J. R. Singler, and Y. Zhang, *A New HDG Method for Dirichlet Boundary Control of Convection Diffusion PDEs* I: *High Regularity*, arXiv:1801.01461, 2018.

[33] W. Hu, J. Shen, J. R. Singler, Y. Zhang, and X. Zheng, *A Superconvergent Hybridizable Discontinuous Galerkin Method for Dirichlet Boundary Control of Elliptic PDEs*, arXiv:1712.02931, 2017.

[34] Y. Jin, *Local analysis of HDG methods for convection-dominated diffusion problems*, J. Comput. Appl. Math., 304 (2016), pp. 57–72, https://doi.org/10.1016/j.cam.2016.03.014.

[35] C. Johnson, A. H. Schatz, and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, Math. Comp., 49 (1987), pp. 25–38, https://doi.org/10.2307/2008248.

[36] J. Liangyue, Y. Xu, and J. Ryan, *Accuracy-enhancement of discontinuous Galerkin solutions for convection-diffusion equations in multiple-dimensions*, Math. Comp., 81 (2012), pp. 1929–1950, https://doi.org/10.1090/S0025-5718-2012-02586-5.

[37] D. Leykekhman and M. Heinkenschloss, *Local error analysis of discontinuous Galerkin methods for advection-dominated elliptic linear-quadratic optimal control problems*, SIAM J. Numer. Anal., 50 (2012), pp. 2012–2038, https://doi.org/10.1137/110826953.

[38] B. Li and X. Xie, *Analysis of a family of HDG methods for second order elliptic problems*, J. Comput. Appl. Math., 307 (2016), pp. 37–51, https://doi.org/10.1016/j.cam.2016.04.027.

[39] M. Mateos and I. Neitzel, *Dirichlet control of elliptic state constrained problems*, Comput. Optim. Appl., 63 (2016), pp. 825–853, https://doi.org/10.1007/s10589-015-9784-y.

[40] S. May, R. Rannacher, and B. Vexler, *Error analysis for a finite element approximation of elliptic Dirichlet boundary control problems*, SIAM J. Control Optim., 51 (2013), pp. 2585–2611, https://doi.org/10.1137/080735734.

[41] K. Niijima, *Pointwise error estimates for a streamline diffusion finite element scheme*, Numer. Math., 56 (1990), pp. 707–719, https://doi.org/10.1007/BF01405198.

[42] W. Qiu and K. Shi, *An HDG method for convection diffusion equation*, J. Sci. Comput., 66 (2016), pp. 346–357, https://doi.org/10.1007/s10915-015-0024-5.

[43] W. Qiu and K. Shi, *A superconvergent HDG method for the incompressible Navier-Stokes equations on general polyhedral meshes*, IMA J. Numer. Anal., 36 (2016), pp. 1943–1967, https://doi.org/10.1093/imanum/drv067.