A Class of Embedded DG Methods for Dirichlet Boundary Control of Convection Diffusion PDEs

Gang Chen1,2 · Guosheng Fu3 · John R. Singler4 · Yangwen Zhang4,5

Received: 21 November 2018 / Revised: 17 May 2019 / Accepted: 22 August 2019 / Published online: 4 September 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
We investigated a hybridizable discontinuous Galerkin (HDG) method for a convection diffusion Dirichlet boundary control problem in our earlier work (Gong et al. SIAM J Numer Anal 56(4):2262–2287, 2018) and obtained an optimal convergence rate for the control under some assumptions on the desired state and the domain. In this work, we obtain the same convergence rate for the control using a class of embedded DG methods proposed by Nguyen et al. (J Comput Phys 302:674–692, 2015) for simulating fluid flows. Since the global system for embedded DG methods uses continuous elements, the number of degrees of freedom for the embedded DG methods are smaller than the HDG method, which uses discontinuous elements for the global system. Moreover, we introduce a new simpler numerical analysis technique to handle low regularity solutions of the boundary control problem. We present some numerical experiments to confirm our theoretical results.

Keywords Dirichlet boundary control · Elliptic convection diffusion equations · Embedded discontinuous Galerkin (EDG) method · Interior embedded discontinuous Galerkin (IEDG) method · Error analysis

Mathematics Subject Classification 65N30 · 49M25

✉ Yangwen Zhang
ywzfg4@mst.edu; ywzhangf@udel.edu
Gang Chen
cglwdm@scu.edu.cn
Guosheng Fu
guosheng_fu@brown.edu
John R. Singler
singlerj@mst.edu

1 School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China
2 College of Mathematics, Sichuan University, Chengdu, China
3 Division of Applied Mathematics, Brown University, Providence, RI, USA
4 Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO, USA
5 Present Address: Department of Mathematical Sciences, University of Delaware, Newark, DE, USA
1 Introduction

We study the following Dirichlet boundary control problem: Minimize the cost functional

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

subject to

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

$$y = u \quad \text{on} \ \partial \Omega, \quad (2)$$

where $\Omega \subset \mathbb{R}^d$ ($d \geq 2$) is a Lipschitz polyhedral domain with boundary $\Gamma = \partial \Omega$. In the 2D case, the optimal control problem (1)–(2) has been proven in [24] to be equivalent to the following optimality system

$$-\varepsilon \Delta z - \nabla \cdot (\beta z) = y - y_d \quad \text{in} \ \Omega, \quad (3c)$$

$$z = 0 \quad \text{on} \ \partial \Omega, \quad (3d)$$

$$\varepsilon \partial_n z - \gamma u = 0 \quad \text{on} \ \partial \Omega. \quad (3e)$$

Dirichlet boundary control plays an important role in many applications; see, e.g., [20, 21, 23, 27, 34] for flow control problems. Approximating the solution of a Dirichlet boundary control problem can be very difficult since solutions frequently have low regularity. Rigorous convergence results have only recently been obtained for Dirichlet boundary control for the Poisson equation using the continuous Galerkin (CG) method [2, 6, 7, 17, 28–30] and a mixed finite element method [19]. A potential weakness of the CG method is that the control and state spaces are coupled: the control space is the trace of the state space. A mixed method allows the control and state spaces to be decoupled, which provides greater flexibility compared to the CG method; however, the degrees of freedom are larger than the CG scheme. It is worth mentioning that Apel et al. in [2] is the first work to obtain a superlinear convergence rate for the control on convex polygonal domains if one uses a superconvergence mesh.

Recently, researchers have investigated discontinuous Galerkin (DG) methods for Dirichlet boundary control problems. We used a hybridizable discontinuous Galerkin (HDG) method for the Poisson equation in [25], and obtained a superlinear convergence rate for the control without using a special mesh or a higher order element. More recently, convection diffusion Dirichlet boundary control problems have gained more and more attention. Benner et al. in [3] used a local discontinuous Galerkin (LDG) method to obtain a sublinear convergence rate for the control. We considered an HDG method and proved optimal superlinear convergence rate for the control in [24] if the regularity of the solution is high, i.e., $y \in H^{1+s}(\Omega)$ with $s \geq 1/2$. To overcome the difficulty for the low regularity case ($0 \leq s < 1/2$), we utilized a special projection operator to get an optimal superlinear convergence rate in [18]; the numerical analysis was more complicated than in [24]. Furthermore, in contrast to [25], we obtained an optimal superlinear convergence rate for the control by using a discontinuous higher order (quadratic) element.

Although the degrees of freedom of the HDG method are significantly reduced compared to standard mixed methods, DG methods and LDG methods, they are still larger than the degrees of freedom of the CG method. In this work, we use embedded DG (EDG) and interior EDG (IEDG) methods to approximate the solution of the Dirichlet boundary control problem.
The EDG and IEDG methods are obtained from the HDG methods, and the global systems both use the same continuous elements; this reduces the number of degrees of freedom considerably. To approximate the control, we use continuous element in the EDG method, and discontinuous elements in the IEDG method. Although the degrees of freedom of IEDG is slightly larger than the EDG method, the IEDG method provides greater flexibility for boundary control problems: we can use different finite element spaces for the control and the state. One possible benefit of the greater flexibility of the IEDG method is that discontinuous elements for the control may be better for more complicated problems (such as convection dominated problems) with sharp changes in the solution. For more details about the EDG and IEDG methods; see Sect. 2.1.

Cockburn et al. in [14] gave a rigorous error analysis of one EDG method for the Poisson equation. Recently, Zhang et al. in [35] proposed a new optimal EDG method for the Poisson equation. Moreover, we used these two EDG methods to approximate the solutions of distributed control problems for the Poisson equation [37] and a convection diffusion equation [36], respectively. However, the techniques in the previous EDG works are not applicable for the Dirichlet boundary control problem since the regularity of the solution may be low. Instead of introducing a special projection as in [18], we use an improved trace inequality from [5] to deal with the low regularity solution. We improve the existing EDG error analysis by dealing with the case of low regularity solutions; also this is the first work to give a rigorous error analysis for the IEDG method. Moreover, in Sect. 3 we prove the same convergence rates for the EDG and IEDG methods that we obtained for HDG methods in [18, 24]. We present numerical results in Sect. 4 for both diffusion dominated and convection dominated problems. Our experiments indicate that both methods work well for both cases; in addition, the IEDG method does a good job at computing sharp changes in the optimal control in the difficult convection dominated case.

2 Background: Regularity and EDG Formulation

Throughout, the standard notation $H^m(\Omega)$ is used for Sobolev spaces on $\Omega$, and we let $\|\cdot\|_{m,\Omega}$ and $|\cdot|_{m,\Omega}$ denote the Sobolev norm and seminorm. We omit the index $m$ when $m = 0$ and the domain $\Omega$ if it will not cause confusion. Also, set $H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$ and $H(\text{div}, \Omega) = \{v \in [L^2(\Omega)]^d, \nabla \cdot v \in L^2(\Omega)\}$. We denote $(\cdot, \cdot)_K$ and $(\cdot, \cdot)_E$ the standard $L^2$-inner products on the domains $K \subset \mathbb{R}^d$ and $E \subset \mathbb{R}^{d-1}$.

Let $\omega (1 < \pi / \omega \leq 3)$ denote the largest interior angle of the domain $\Omega$, i.e., $\Omega$ is a convex polygonal domain. Moreover, we assume $\beta$ satisfies

$$\beta \in [L^\infty(\Omega)]^d, \quad \nabla \cdot \beta \in L^\infty(\Omega), \quad \nabla \cdot \beta \leq 0, \quad \nabla \nabla \cdot \beta \in [L^2(\Omega)]^d.$$  

The mixed weak form of the formal optimality system (3a)–(3e) is

$$\varepsilon^{-1}(q, r)_\Omega - (y, \nabla \cdot r)_\Omega + (u, r \cdot n)_\Gamma = 0, \quad (5a)$$

$$\nabla \cdot (q + \beta y), w)_\Omega - (y \nabla \cdot \beta, w)_\Omega = (f, w)_\Omega, \quad (5b)$$

$$\varepsilon^{-1}(p, r)_\Omega - (z, \nabla \cdot r)_\Omega = 0, \quad (5c)$$

$$\nabla \cdot (p - \beta z), w)_\Omega = (y - y_d, w)_\Omega, \quad (5d)$$

$$\langle y u + p \cdot n, \mu \rangle_\Gamma = 0 \quad (5e)$$

for all $(r, w, \mu) \in H(\text{div}, \Omega) \times L^2(\Omega) \times L^2(\Gamma)$.

The following well-posedness and regularity result is found in [24].
Theorem 2.1 If \( f = 0 \) and \( y_d \in H^{1+r} (\Omega) \) for some \( 0 \leq r < 1 \), then the optimal control problem (1)–(2) has a unique solution \( u \in L^2 (\Gamma) \) and \( u \) is uniquely determined by the optimality system (5a)–(5e). Moreover, for any \( s > 0 \) satisfying \( s \leq 1/2 + r \) and \( s < \min (3/2, \pi/\omega - 1/2) \), we have \( u \in H^s (\Gamma) \) and

\[
(q, p, y, z) \in \left[ H^{s-\frac{1}{2}}(\Omega) \right]^d \cap H(\text{div}, \Omega) \times \left[ H^{s+\frac{1}{2}}(\Omega) \right]^d \times H^{s+\frac{1}{2}}(\Omega) \times H^{s+\frac{3}{2}}(\Omega).
\]

We note that the case of \( f \neq 0 \) can be handled by the technique in [1, p. 3623]. Theorem 2.1 implies that if \( y_d \in H^{1+r} (\Omega) \) for some \( r \in (1/2, 1) \), and \( \pi/3 < \omega < 2\pi/3 \), then \( u \in H^{1+r} (\Gamma) \) for some \( r_u \in (1, 3/2) \), we called this the high regularity case in [24]. In this scenario, \( q \in H^s (\Omega) \) with \( r_q > 1/2 \), which guarantees \( q \) has a \( L^2 \) boundary trace. We used this property to give a convergence analysis of HDG methods in [24,25].

However, if \( r \in [0, 1/2) \) or \( 2\pi/3 \leq \omega < \pi \), then we are in the low regularity case, i.e., \( u \in H^{1+r} (\Gamma) \) for some \( r_u \in [1/2, 1) \), and \( \| q \|_{\partial \Omega} \) is not well-defined. The numerical analysis is more difficult in this case; see [18] for an HDG method in the low regularity case.

2.1 A Class of Embedded DG Formulations

To better describe the class of Embedded DG (EDG) methods, we first give some notation.

Let \( \mathcal{T}_h \) be a conforming, quasi-uniform triangulation of \( \Omega \). We denote by \( \partial \mathcal{T}_h \) the set \{\( \partial K : K \in \mathcal{T}_h \)\}. For \( K \in \mathcal{T}_h \), let \( e = \partial K \cap \Gamma \) denote the boundary face of \( K \) if the \( d - 1 \) Lebesgue measure of \( e \) is non-zero. For two elements \( K_1, K_2 \in \mathcal{T}_h \), let \( e = \partial K_1 \cap \partial K_2 \) denote the interior face between \( K_1 \) and \( K_2 \) if the \( d - 1 \) Lebesgue measure of \( e \) is non-zero. Let \( \mathcal{E}_h^o \) and \( \mathcal{E}_h^\partial \) denote the sets of interior and boundary faces, respectively. We denote by \( \mathcal{E}_h \) the union of \( \mathcal{E}_h^o \) and \( \mathcal{E}_h^\partial \). Finally, we introduce

\[
(w, v)_{\mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} (w, v)_K,
\]

\[
\langle \xi, \rho \rangle_{\partial \mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} \langle \xi, \rho \rangle_{\partial K}.
\]

HDG methods were proposed by Cockburn et al. in [11] as an improvement of traditional discontinuous Galerkin (DG) methods and have many applications; see, e.g., [8,9,12,13,15,16,32]. HDG methods are based on mixed formulations and introduce a new variable to approximate the trace of the scalar variable along the element boundary. To approximate the flux variable and solution, we use the discontinuous finite element spaces \( V_h \) and \( W_h \):

\[
V_h := \{ v \in [L^2(\Omega)]^d : v|_K \in [\mathcal{P}^k(K)]^d, \forall K \in \mathcal{T}_h \},
\]

\[
W_h := \{ w \in L^2(\Omega) : w|_K \in \mathcal{P}^\ell(K), \forall K \in \mathcal{T}_h \},
\]

where \( \mathcal{P}^k(K) \) denotes the set of polynomials of degree at most \( k \) on a domain \( K \). HDG methods use the discontinuous finite element spaces to express the approximate flux and solution in an element-by-element fashion in terms of numerical traces of the scalar variable. Then the globally coupled system only involves the numerical trace. The high number of globally coupled degrees of freedom is significantly reduced compared to other DG methods and standard mixed methods.

For the HDG methods, we use the following discontinuous finite element space to approximate the numerical trace:

\[
M_h^\text{HDG} := \{ \mu \in L^2(\mathcal{E}_h) : \mu|_e \in \mathcal{P}^m(e), \forall e \in \mathcal{E}_h \}.
\]
Note that $M_{h}^{\text{HDG}}$ consists of functions which are discontinuous at the border of the faces. Embedded discontinuous Galerkin (EDG) methods, which were originally proposed in [22], are obtained from HDG methods by replacing the discontinuous finite element space for the numerical traces with a continuous space, i.e.,

$$
M_{h}^{\text{EDG}} := \{ \mu \in C^{0}(\mathcal{E}_{h}) : \mu|_{e} \in \mathcal{P}^{m}(e), \forall e \in \mathcal{E}_{h} \}.
$$

Hence, the number of degrees of freedom for the EDG method are much smaller than the HDG method, and also the same with the CG method (after static condensation). The interior embedded discontinuous Galerkin (IEDG) method was proposed and investigated for convection dominated flow problems in [31]. The IEDG method is obtained by a simple change to the space of the numerical trace from the HDG and EDG methods; specifically,

$$
M_{h}^{\text{IEDG}} := \{ \mu \in L^{2}(\mathcal{E}_{h}) : \mu|_{e} \in \mathcal{P}^{m}(e), \forall e \in \mathcal{E}_{h}, \text{ and } \mu|_{\mathcal{E}_{h}^{0}} \in C^{0}(\mathcal{E}_{h}^{0}) \}.
$$

The functions in $M_{h}^{\text{IEDG}}$ are only continuous on the union of interior edges and are discontinuous on the union of the boundary edges. This simple change has many benefits even for pure PDE simulations; see [31] for details. Compared to the EDG methods, the IEDG methods have a great potential for boundary control problems since they allow us to choose different spaces for the state and the control, as discussed in the introduction.

In this paper, we perform a numerical analysis for both EDG and IEDG methods for the convection diffusion Dirichlet boundary control problem. To unify the analysis, we omit the superscripts EDG and IEDG on the space $M_{h}^{\text{EDG}}$ and $M_{h}^{\text{IEDG}}$, respectively. We choose the following finite element spaces:

$$
V_{h} := \{ v \in [L^{2}(\Omega)]^{d} : v|_{K} \in [\mathcal{P}^{k}(K)]^{d}, \forall K \in \mathcal{T}_{h} \},
$$

$$
W_{h} := \{ w \in L^{2}(\Omega) : w|_{K} \in \mathcal{P}^{k+1}(K), \forall K \in \mathcal{T}_{h} \},
$$

$$
M_{h} := \{ \mu \in L^{2}(\mathcal{E}_{h}) : \mu|_{e} \in \mathcal{P}^{k+1}(e), \forall e \in \mathcal{E}_{h}, \text{ and } \mu|_{\mathcal{E}_{h}^{0}} \in C^{0}(\mathcal{E}_{h}^{0}) \}.
$$

Let $M_{h}(\partial)$ be the spaces defined similarly to $M_{h}$ with $\mathcal{E}_{h}$ replaced by the set of interior edges $\mathcal{E}_{h}^{0}$ and the set of boundary edges $\mathcal{E}_{h}^{1}$, respectively. The functions in $M_{h}(\partial)$ are continuous for both the EDG method and IEDG method, while the functions in $M_{h}(\partial)$ are continuous across the boundary edges for the EDG method and discontinuous for the IEDG method. In addition, for any function $w \in W_{h}$ and $r \in V_{h}$, we use $\nabla w$ and $\nabla \cdot r$ to denote the piecewise gradient and divergence on each element $K \in \mathcal{T}_{h}$, respectively.

Below, we consider the EDG and IEDG methods simultaneously; the choice of $M_{h}(\partial)$ determines the method as indicated above. The EDG (or IEDG) method seeks approximate fluxes $q_{h}$, $p_{h} \in V_{h}$, states $y_{h}, z_{h} \in W_{h}$, interior element boundary traces $\tilde{y}_{h}, \tilde{z}_{h} \in M_{h}(\partial)$, and control $u_{h} \in M_{h}(\partial)$ satisfying

$$
\varepsilon^{-1}(q_{h}, r_{1})_{\mathcal{T}_{h}} - (y_{h}, \nabla \cdot r_{1})_{\mathcal{T}_{h}} + (\tilde{y}_{h}, r_{1})_{\partial \mathcal{T}_{h} \setminus \mathcal{E}_{h}^{a}} + (u_{h}, r_{1})_{\mathcal{E}_{h}^{a}} = 0, \quad (6a)
$$

$$
- (q_{h} + \beta y_{h}, \nabla w_{1})_{\mathcal{T}_{h}} - (y_{h} \nabla \cdot \beta, w_{1})_{\mathcal{T}_{h}} + (\tilde{q}_{h} \cdot n, w_{1})_{\partial \mathcal{T}_{h}} + (\beta \cdot n u_{h}, w_{1})_{\mathcal{E}_{h}^{a}} = (f, w_{1})_{\mathcal{T}_{h}}, \quad (6b)
$$

for all $(r_{1}, w_{1}) \in V_{h} \times W_{h}$,

$$
\varepsilon^{-1}(p_{h}, r_{2})_{\mathcal{T}_{h}} - (z_{h}, \nabla \cdot r_{2})_{\mathcal{T}_{h}} + (\tilde{z}_{h}, r_{2})_{\partial \mathcal{T}_{h} \setminus \mathcal{E}_{h}^{a}} = 0, \quad (6c)
$$

$$
- (p_{h} - \beta z_{h}, \nabla w_{2})_{\mathcal{T}_{h}} + (\tilde{p}_{h} \cdot n, w_{2})_{\partial \mathcal{T}_{h}} - (\beta \cdot n \tilde{z}_{h}, w_{2})_{\partial \mathcal{T}_{h} \setminus \mathcal{E}_{h}^{a}} - (y_{h}, w)_{\mathcal{T}_{h}} = -(y_{d}, w_{2})_{\mathcal{T}_{h}}, \quad (6d)
$$

\text{Springer}
for all \((r_2, w_2) \in V_h \times W_h\),

\[
\langle \hat{q}_h \cdot n, \mu_1 \rangle_{\partial T_h \setminus \epsilon} = 0, \quad (6e)
\]
\[
\langle \hat{p}_h \cdot n, \mu_2 \rangle_{\partial T_h \setminus \epsilon} = 0, \quad (6f)
\]

for all \(\mu_1, \mu_2 \in M_h(o)\), and the optimality condition

\[
\gamma \langle u_h, \mu_3 \rangle_{E} + \langle \hat{p}_h \cdot n, \mu_3 \rangle_{E} = 0, \quad (6g)
\]

for all \(\mu_3 \in M_h(\partial)\). The numerical traces on \(\partial T_h\) are defined as

\[
\hat{q}_h \cdot n = q_h \cdot n + (h^{-1} + \tau_1)(y_h - \hat{y}_h) \quad \text{on } \partial T_h \setminus \epsilon, \quad (6h)
\]
\[
\hat{q}_h \cdot n = q_h \cdot n + (h^{-1} + \tau_1)(y_h - u_h) \quad \text{on } \epsilon, \quad (6i)
\]
\[
\hat{p}_h \cdot n = p_h \cdot n + (h^{-1} + \tau_2)(z_h - \hat{z}_h) \quad \text{on } \partial T_h \setminus \epsilon, \quad (6j)
\]
\[
\hat{p}_h \cdot n = p_h \cdot n + (h^{-1} + \tau_2)z_h \quad \text{on } \epsilon, \quad (6k)
\]

where \(\tau_1\) and \(\tau_2\) are positive stabilization functions defined on \(\partial T_h\) that satisfy

\[
\tau_2 = \tau_1 - \beta \cdot n. \quad (7)
\]

The condition (7) for the stabilization functions \(\tau_1\) and \(\tau_2\) has now been used in a number of works; see, e.g., [18,24] for convection diffusion Dirichlet boundary control problems and [10,26,36] for convection diffusion distributed optimal control problems. This condition causes the optimize-then-discretize and discretize-then-optimize EDG/HDG approaches to the control problem to produce equivalent results; see [36] for details concerning an EDG method for a distributed convection diffusion optimal control problem. Our implementation of the EDG and IEDG methods is similar to the HDG implementation for a Poisson Dirichlet boundary control problem described in our earlier work [25].

### 3 Error Analysis

Next, we provide a convergence analysis of the above EDG and IEDG methods for the convection diffusion Dirichlet boundary control problem in both high regularity and low regularity cases. For the high regularity case, tools from the analysis technique in [36] for a convection diffusion distributed control problem can be modified to apply to the Dirichlet boundary control problem. For the low regularity case, we introduced a special projection operator in our earlier HDG work [18] to avoid the quantity \(\|q \cdot n\|_{\partial T_h}\) in the analysis; however, this complicated the analysis. In this work, we use an improved inverse inequality from [5], and simplify the error analysis for the low regularity case. It is worth mentioning that part of our analysis (step 1 to step 3 in Sect. 3.3) improves the existing EDG error analysis by dealing with the case of low regularity solutions. In this work, we only perform an error analysis for the diffusion dominated case; i.e., in this section, we assume \(\varepsilon = O(1)\). The generic constant \(C\) may depend on the data of the problem but is independent of \(h\) and may change from line to line.
3.1 Assumptions and Main Result

We assume the solution of the optimality system (5a)–(5e) has the following regularity properties:

\[
q \in [H^q(\Omega)]^d \cap H(\text{div}, \Omega), \quad p \in [H^r(\Omega)]^d, \quad y \in H^r(\Omega), \quad z \in H^r(\Omega),
\]

\[
r_q > 0, \quad r_p > 1, \quad r_y > 1, \quad r_z > 2.
\]  

(8a)

In the 2D case, Theorem 2.1 guarantees this regularity condition is satisfied.

We now state our main convergence result.

Theorem 3.1 Let

\[
s_q = \min\{r_q, 1\}, \quad s_y = \min\{r_y, k + 2\},
\]

\[
s_p = \min\{r_p, k + 1\}, \quad s_z = \min\{r_z, k + 2\}.
\]  

(9)

we have

\[
\begin{align*}
\|u - u_h\|_{\mathcal{E}_0^q} & \leq C \left( h^{s_q + \frac{1}{2}} \|q\|_{s_q} + h^{s_y - \frac{1}{2}} \|y\|_{s_y} + h^{s_p - \frac{1}{2}} \|p\|_{s_p} + h^{s_z - \frac{3}{2}} \|z\|_{s_z} \right), \\
\|y - y_h\|_{\mathcal{T}_h} & \leq C \left( h^{s_q + \frac{1}{2}} \|q\|_{s_q} + h^{s_y - \frac{1}{2}} \|y\|_{s_y} + h^{s_p - \frac{1}{2}} \|p\|_{s_p} + h^{s_z - \frac{3}{2}} \|z\|_{s_z} \right), \\
\|p - p_h\|_{\mathcal{T}_h} & \leq C \left( h^{s_q + \frac{1}{2}} \|q\|_{s_q} + h^{s_y - \frac{1}{2}} \|y\|_{s_y} + h^{s_p - \frac{1}{2}} \|p\|_{s_p} + h^{s_z - \frac{3}{2}} \|z\|_{s_z} \right), \\
\|z - z_h\|_{\mathcal{T}_h} & \leq C \left( h^{s_q + \frac{1}{2}} \|q\|_{s_q} + h^{s_y - \frac{1}{2}} \|y\|_{s_y} + h^{s_p - \frac{1}{2}} \|p\|_{s_p} + h^{s_z - \frac{3}{2}} \|z\|_{s_z} \right).
\end{align*}
\]

If \( k \geq 1 \), then

\[
\|q - q_h\|_{\mathcal{T}_h} \leq C \left( h^{s_q} \|q\|_{s_q} + h^{s_y - 1} \|y\|_{s_y} + h^{s_p - 1} \|p\|_{s_p} + h^{s_z - 2} \|z\|_{s_z} \right).
\]

Specializing to the 2D case gives the following result:

Corollary 3.2 Suppose \( d = 2 \), \( f = 0 \) and \( y_d \in H^r(\Omega) \) for some \( t^* \in (0, 1) \). Let \( \pi/3 \leq \omega < \pi \) be the largest interior angle of \( \Gamma \), and let \( r > 0 \) satisfy

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

If \( k = 1 \), then

\[
\begin{align*}
\|u - u_h\|_{\mathcal{E}_0^q} & \leq C h^r \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^{r+1/2}} + \|z\|_{H^{r+3/2}} \right), \\
\|y - y_h\|_{\mathcal{T}_h} & \leq C h^r \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^{r+1/2}} + \|z\|_{H^{r+3/2}} \right), \\
\|p - p_h\|_{\mathcal{T}_h} & \leq C h^r \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^{r+1/2}} + \|z\|_{H^{r+3/2}} \right), \\
\|z - z_h\|_{\mathcal{T}_h} & \leq C h^r \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^{r+1/2}} + \|z\|_{H^{r+3/2}} \right).
\end{align*}
\]

If in addition \( r > 1/2 \), then

\[
\|q - q_h\|_{\mathcal{T}_h} \leq C h^{-1/2} \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^{r+1/2}} + \|z\|_{H^{r+3/2}} \right).
\]

Furthermore, if \( k = 0 \) then

\[
\begin{align*}
\|u - u_h\|_{\mathcal{E}_0^q} & \leq C h^{1/2} \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^1} + \|z\|_{H^2} \right), \\
\|y - y_h\|_{\mathcal{T}_h} & \leq C h^{1/2} \left( \|q\|_{H^{r-1/2}} + \|y\|_{H^{r+1/2}} + \|p\|_{H^1} + \|z\|_{H^2} \right),
\end{align*}
\]
the Scott-Zhang interpolation operator

Therefore, the standard Lagrange interpolation operator is not applicable; hence we utilize

Moreover, we use the following well-known bounds:

Remark 3.3 As in [18,24], when \( k = 1 \) the convergence rates are optimal for the control and
the flux \( q \) and suboptimal for the other variables. Compared to the HDG method used in
[18,24], we obtain the same convergence rates for the EDG and IEDG methods.

We also note that our current analysis approach is not able to deal with the case \( r^* = 0 \),
i.e., \( y_d \) is only in \( L^2(\Omega) \) without any additional regularity. In this case, \( q \) is only guaranteed to
be in \( H(\text{div}) \) without any additional regularity, and the trace inequality below in Lemma 3.7
is not applicable. It may be possible to treat this case using some other approach; we leave
this to be considered elsewhere.

3.2 Preliminary Material

We introduce the standard \( L^2 \)-orthogonal projection operators \( \Pi_h^k : [L^2(K)]^d \to [P^k(K)]^d \)
and \( \Pi_h^{k+1} : L^2(K) \to P^{k+1}(K) \), which satisfy

Moreover, we use the following well-known bounds:

where \( s_q \) and \( s_y \) are defined in Theorem 3.1. We have the same projection error bounds for
\( p \) and \( z \).

Since we only assume \( y \in H^s(\Omega) \) with \( r_y > 1 \), certain components of the solution may
not be continuous; for example, we cannot guarantee \( y \) is continuous on \( \Omega \) when \( d = 3 \).
Therefore, the standard Lagrange interpolation operator is not applicable; hence we utilize
the Scott-Zhang interpolation operator \( \tilde{I}_h^{k+1} : H^1(\Omega) \to \tilde{W}_h \) from [33], where

The following bound is found in [33, Theorem 4.1]:

By an inverse inequality, a trace inequality and Eq. (12) we obtain

Next, for any \( (q_h, y_h, \tilde{y}_h^o, r_1, w_1, \mu_1) \in [V_h \times W_h \times M_h(o)]^2 \) and \( (p_h, z_h, \tilde{z}_h^o, r_2, w_2, \mu_2) \in [V_h \times W_h \times M_h(o)]^2 \), define the operators \( B_1 \) and \( B_2 \) by

\( e_1 \) Springer
Lemma 3.6 There exists a unique solution of the discrete system (16).

Next, we present three basic but fundamental results. The proofs follow similar arguments in [18,24,25] and are omitted.

Lemma 3.4 For any \( (v_h, w_h, \mu_h) \in V_h \times W_h \times M_h(\partial) \), we have

\[
\mathcal{B}_1(v_h, w_h, \mu_h; v_h, w_h, \mu_h) = \varepsilon^{-1}(v_h, v_h)_T + \left( \left( h^{-1} + \tau_1 - \frac{1}{2} \beta \cdot n \right) (w_h - \mu_h), w_h - \mu_h \right)_{\partial T_h \setminus \mathcal{E}_h^\partial} - \frac{1}{2} (\nabla \cdot w_h)_T + \left( \left( h^{-1} + \tau_1 - \frac{1}{2} \beta \cdot n \right) w_h, w_h \right)_{\mathcal{E}_h^\partial},
\]

\[
\mathcal{B}_2(v_h, w_h, \mu_h; v_h, w_h, \mu_h) = \varepsilon^{-1}(v_h, v_h)_T + \left( \left( h^{-1} + \tau_2 + \frac{1}{2} \beta \cdot n \right) (w_h - \mu_h), w_h - \mu_h \right)_{\partial T_h \setminus \mathcal{E}_h^\partial} - \frac{1}{2} (\nabla \cdot w_h)_T + \left( \left( h^{-1} + \tau_2 + \frac{1}{2} \beta \cdot n \right) w_h, w_h \right)_{\mathcal{E}_h^\partial}.
\]

Lemma 3.5 For any \( (v_1, v_2, w_1, w_2, \mu_1, \mu_2) \in V_h \times V_h \times W_h \times M_h(\partial) \), we have

\[
\mathcal{B}_1(v_1, w_1, \mu_1; v_2, -w_2, -\mu_2) + \mathcal{B}_2(v_2, w_2, \mu_2; -v_1, w_1, \mu_1) = 0.
\]

Proposition 3.6 There exists a unique solution of the discrete system (16).

Next, we introduce the improved trace inequality.

Lemma 3.7 [5, Lemma 2.4] Let \( E \) be a face of \( K \in T_h \). If \( \mathbf{q} \in [H^s_q(\Omega)]^d \cap H(\text{div}, \Omega) \) with \( s_q > 0 \), then for all \( \mu \in \mathcal{P}^{k+1}(E) \), we have

\[
\langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_E \leq C h^{-1/2} \| \mu \|_E (\| \mathbf{q} \|_K + h \| \nabla \cdot \mathbf{q} \|_K).
\]
3.3 Proof of Theorem 3.1

To prove Theorem 3.1, we follow the strategy in [25] and split the proof into seven steps. We consider the following auxiliary problem: find

\[(q_h(u), p_h(u), y_h(u), z_h(u), \tilde{z}_h^0(u), \tilde{z}_h^0(u)) \in V_h \times V_h \times W_h \times M_h(x) \times M_h(x)\]

such that

\[
\begin{aligned}
B_1(q_h(u), y_h(u), \tilde{z}_h^0(u); r_1, w_1, \mu_1) &= -\langle u, r_1 \cdot n - (h^{-1} + \tau_1 - \beta \cdot n) w_1 \rangle e_h^0 \\
&+ (f, w_1)_{T_h}, \tag{18a}
\end{aligned}
\]

\[
\begin{aligned}
B_2(p_h(u), z_h(u), \tilde{z}_h^0(u); r_2, w_2, \mu_2) &= (y_h(u) - y_d, w_2)_{T_h} \tag{18b}
\end{aligned}
\]

for all \((r_1, r_2, w_1, w_2, \mu_1, \mu_2) \in V_h \times V_h \times W_h \times M_h(x) \times M_h(x)\). We begin by bounding the error between the solutions of the auxiliary problem (18) and the mixed form (5a)–(5d) of the optimality system.

3.3.1 Step 1: The Error Equation for Part 1 of The Auxiliary Problem (18a)

**Lemma 3.8** For all \((r_1, w_1, \mu_1) \in V_h \times W_h \times M_h(x)\), we have

\[
\begin{aligned}
B_1(\Pi_h^0 q, \Pi_h^{k+1} y, \tilde{J}_h^{k+1} y, r_1, w_1, \mu_1) &= \langle u, (h^{-1} + \tau_1 - \beta \cdot n) w_1 - r_1 \cdot n \rangle e_h^0 \\
&+ (\tilde{J}_h^{k+1} y - y, r_1 \cdot n)_{\partial T_h \setminus e_h^0} - (\Pi_h^0 q - q, \nabla w_1)_{T_h} \\
&+ (\beta (y - \Pi_h^{k+1} y), \nabla w_1)_{T_h} + (\nabla \cdot \beta (y - \Pi_h^{k+1} y), w_1)_{T_h} \\
&+ ((h^{-1} + \tau_1) (\Pi_h^{k+1} y - \tilde{J}_h^{k+1} y), w_1 - \mu_1)_{\partial T_h \setminus e_h^0} \\
&+ ((\Pi_h^0 q - q) \cdot n, w_1)_{e_h^0}.
\end{aligned}
\]

**Proof** Using the definition of \(B_1\) in (14) gives

\[
\begin{aligned}
B_1(\Pi_h^0 q, \Pi_h^{k+1} y, \tilde{J}_h^{k+1} y, r_1, w_1, \mu_1) &= \langle (\Pi_h^{k+1} y - \tilde{J}_h^{k+1} y, \nabla \cdot r_1)_{T_h} + (\tilde{J}_h^{k+1} y, r_1 \cdot n)_{\partial T_h \setminus e_h^0} \\
&+ (\nabla \cdot \Pi_h^0 q, w_1)_{T_h} - (\beta \Pi_h^{k+1} y, \nabla w_1)_{T_h} - (\nabla \cdot \beta \Pi_h^{k+1} y, w_1)_{T_h} \\
&+ ((h^{-1} + \tau_1) \Pi_h^{k+1} y, w_1)_{\partial T_h \setminus e_h^0} - (\Pi_h^0 q \cdot n + (h^{-1} + \tau_1) (\Pi_h^{k+1} y - \tilde{J}_h^{k+1} y), \mu_1)_{\partial T_h \setminus e_h^0}.
\end{aligned}
\]

Using properties of the \(L^2\) projections (10) gives

\[
\begin{aligned}
B_1(\Pi_h^0 q, \Pi_h^{k+1} y, \tilde{J}_h^{k+1} y, r_1, w_1, \mu_1) &= \langle u, r_1 \cdot n \rangle_{\partial T_h \setminus e_h^0} - (y, \nabla \cdot r_1)_{T_h} + \langle y, r_1 \cdot n \rangle_{\partial T_h \setminus e_h^0} \\
&- \langle \Pi_h^0 q - r_1 \rangle_{T_h} + (\tilde{J}_h^{k+1} y - y, r_1 \cdot n)_{\partial T_h \setminus e_h^0} \\
&+ (\nabla \cdot q, w_1)_{T_h} + (\nabla \cdot (\Pi_h^0 q - q), w_1)_{T_h} - (\beta y, \nabla w_1)_{T_h}.
\end{aligned}
\]
\[ + (\beta(y - \Pi_h^{k+1} y), \nabla w_1)_T_h - (\nabla \cdot \beta y, w_1)_T_h + (\nabla \cdot \beta(y - \Pi_h^{k+1} y), w_1)_T_h \\
+ ((h^{-1} + \tau_1)(\Pi_h^{k+1} y - y), w_1)_{\partial T_h} - ((h^{-1} + \tau_1)(\Pi_h^{k+1} y - y), w_1)_{\partial T_h} e_h^a \\
+ ((h^{-1} + \tau_1)y, w_1)_\mathcal{E}_h^a + (\beta \cdot n y, w_1)_{\partial T_h} e_h^a + (\beta \cdot n(\Pi_h^{k+1} y - y), w_1)_{\partial T_h} e_h^a \\
- (\Pi_0^q \cdot n + (h^{-1} + \tau_1)(\Pi_h^{k+1} y - \mathcal{I}_h^{k+1} y), \mu_1)_{\partial T_h} e_h^a. \]

The flux \( q \) and state \( y \) satisfy

\[ e^{-1}(q \cdot r_1)_T_h - (y, \nabla \cdot r_1)_T_h + (y, r_1 \cdot n)_{\partial T_h} e_h^a = -(u, r_1 \cdot n) e_h^a, \]

\[ (\nabla \cdot q, w_1)_T_h - (\beta y, \nabla w_1)_T_h - (\nabla \cdot \beta y, w_1)_T_h \\
+ (\beta \cdot n y, w_1)_{\partial T_h} e_h^a = -(\beta \cdot n u, w_1)_e_h^a + (f, w_1)_T_h, \]

\[ (q \cdot n, \mu_1)_{\partial T_h} e_h^a = 0 \]

for all \((r_1, w_1, \mu_1) \in \mathcal{V}_h \times \mathcal{W}_h \times M_h(o)\). This gives

\[ \mathcal{B}_1(\Pi_0^q, \Pi_h^{k+1} y, \mathcal{I}_h^{k+1} y, r_1, w_1, \mu_1) \]

\[ = (u, (h^{-1} + \tau_1 - \beta \cdot n) w_1 - r_1 \cdot n)_{e_h^a} + (f, w_1)_T_h - e^{-1}(q - \Pi_h^0 q, r_1)_T_h \\
+ (\mathcal{I}_h^{k+1} y - y, r_1 \cdot n)_{\partial T_h} e_h^a + (\nabla \cdot (\Pi_0^q - q), w_1)_T_h \\
+ (\beta(y - \Pi_h^{k+1} y), \nabla w_1)_T_h + (\nabla \cdot \beta(y - \Pi_h^{k+1} y), w_1)_T_h \\
+ ((h^{-1} + \tau_1)(\mathcal{I}_h^{k+1} y - \mathcal{I}_h^{k+1} y), w_1 - \mu_1)_{\partial T_h} e_h^a \\
+ ((h^{-1} + \tau_1)(\mathcal{I}_h^{k+1} y - y), w_1)_{e_h^a} + (\beta \cdot n(\mathcal{I}_h^{k+1} y - y), w_1)_{\partial T_h} e_h^a \\
- (\Pi_0^q - q) \cdot n - \mu_1)_{\partial T_h} e_h^a + (\Pi_0^q - q) \cdot n, w_1 - \mu_1)_{\partial T_h} e_h^a \\
+ ((\Pi_0^q - q) \cdot n, w_1)_{e_h^a}. \]

\[ \square \]

**Lemma 3.9** For all \((r_2, w_2, \mu_2) \in \mathcal{V}_h \times \mathcal{W}_h \times M_h(o)\), we have

\[ \mathcal{B}_2(\Pi_h^k p, \Pi_h^{k+1} z, \mathcal{I}_h^{k+1} z, r_2, w_2, \mu_2) \]

\[ = (y - y_d, w_2)_T_h + (\mathcal{I}_h^{k+1} z - z, r_2 \cdot n)_{\partial T_h} e_h^a \\
- (\beta(z - \Pi_h^{k+1} z), \nabla w_2)_T_h + ((h^{-1} + \tau_2)(\mathcal{I}_h^{k+1} z - \mathcal{I}_h^{k+1} z), w_2 - \mu_2)_{\partial T_h} e_h^a \\
+ ((h^{-1} + \tau_2)(\mathcal{I}_h^{k+1} z - z), w_2 - \mu_2)_{e_h^a} - (\beta \cdot n(\mathcal{I}_h^{k+1} z - z), w_2)_{\partial T_h} e_h^a \\
+ ((\Pi_h^k p - p) \cdot n, w_2 - \mu_2)_{\partial T_h} e_h^a + ((\Pi_h^k p - p) \cdot n, w_2)_{e_h^a}. \]
The proof proceeds in the same way as the proof of the above lemma.
Subtracting part 1 of the auxiliary problem (18a) from the equality in Lemma 3.8 gives the following result:

**Lemma 3.10** For $\epsilon^q_h = \Pi_h^0 q - q_h(u), \epsilon^y_h = \Pi_h^{k+1} y - y_h(u), \epsilon_{\tilde{y}}^h = \tilde{\tau}^{k+1} y - \tilde{y}_h(u)$, we have

$$
\mathcal{B}_1(\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h, r_1, w_1, \mu_1) \\
= -\epsilon^{-1}(q - \Pi_h^0 q, q, r_1)_{T_h} + (\tilde{\tau}^{k+1} y - y, r_1 \cdot n)_{\partial T_h \setminus \epsilon^q_h} - (\Pi_h^0 q - q, \nabla w_1)_{T_h} \\
+ (\beta(y - \Pi_h^{k+1} y), \nabla w_1)_{T_h} + (\nabla \beta(y - \Pi_h^{k+1} y), w_1)_{T_h} \\
+ (h^{-1} + \tau_1)(\Pi_h^{k+1} y - \tilde{\tau}^{k+1} y), w_1 - \mu_1)_{\partial T_h \setminus \epsilon^q_h} \\
+ (h^{-1} + \tau_1)(\Pi_h^{k+1} y - y), w_1 - \mu_1)_{\epsilon^q_h} \\
+ (\beta \cdot n(\tilde{\tau}^{k+1} y - y), w_1 - \mu_1)_{\partial T_h \setminus \epsilon^q_h} + ((\Pi_h^0 q - q) \cdot n, w_1 - \mu_1)_{\partial T_h \setminus \epsilon^q_h} \\
+ ((\Pi_h^0 q - q) - n, w_1)_{\epsilon^q_h}
$$

for all $(r_1, w_1, \mu_1) \in V_h \times W_h \times M_h(o)$.

### 3.3.2 Step 2: Estimates for $\epsilon^Q_h$.

**Lemma 3.11** For $(\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h)$ defined in Lemma 3.10, we have

$$
\| \epsilon^q_h \|_{T_h} + h^{-\frac{1}{2}} \| \epsilon^y_h - \epsilon_{\tilde{y}}^h \|_{\partial T_h \setminus \epsilon^q_h} + h^{-\frac{1}{2}} \| \epsilon^y_h \|_{\epsilon^q_h} \\
\leq C \| q - \Pi_h^0 q \|_{T_h} + C h \| \nabla q \|_{T_h} \\
+ C h^{-1/2}(\| \Pi_h^{k+1} y - y \|_{\partial T_h} + \| \tilde{\tau}^{k+1} y - y \|_{\partial T_h}).
$$

**Proof** First, we take $(r_1, w_1, \mu_1) = (\nabla \epsilon^y_h, 0, 0)$ in Lemma 3.10, and by the definition of $\mathcal{B}_1$ in (14) we have

$$
\| \nabla \epsilon^y_h \|_{T_h} \leq C(\| \epsilon^q_h \|_{T_h} + h^{-1/2} \| \epsilon^y_h - \epsilon_{\tilde{y}}^h \|_{\partial T_h \setminus \epsilon^q_h} + h^{-1/2} \| \epsilon^y_h \|_{\epsilon^q_h}) \\
+ C h^{-1/2}(\| \tilde{\tau}^{k+1} y - y \|_{\partial T_h} + \| \Pi_h^0 q \|_{T_h}).
$$

(19)

Next, taking $(r_1, w_1, \mu_1) = (\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h)$ in Lemma 3.4, we get

$$
\mathcal{B}_1(\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h, \epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h) \\
= \epsilon^{-1}(\epsilon^q_h, \epsilon^y_h)_{T_h} + \left(\frac{1}{2} \beta \cdot n, \epsilon^y_h - \epsilon_{\tilde{y}}^h \right)_{\partial T_h \setminus \epsilon^q_h} \\
- \frac{1}{2}(\nabla \beta \epsilon^y_h, \epsilon_{\tilde{y}}^h)_{T_h} + \left(\frac{1}{2} \beta \cdot n \right) \epsilon^y_h, \epsilon_{\tilde{y}}^h \epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h
$$

On the other hand, take $(r_1, w_1, \mu_1) = (\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h)$ in Lemma 3.10 to obtain

$$
\mathcal{B}_1(\epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h, \epsilon^q_h, \epsilon^y_h, \epsilon_{\tilde{y}}^h) \\
= -\epsilon^{-1}(q - \Pi_h^0 q, \epsilon^q_h)_{T_h} + (\tilde{\tau}^{k+1} y - y, \epsilon^q_h \cdot n)_{\partial T_h \setminus \epsilon^q_h} - (\Pi_h^0 q - q, \nabla \epsilon^y_h)_{T_h} \\
+ (\beta(y - \Pi_h^{k+1} y), \nabla \epsilon^y_h)_{T_h} + (\nabla \beta(y - \Pi_h^{k+1} y), \epsilon^y_h)_{T_h}.
$$
For the term $T_1$, the Cauchy-Schwarz inequality gives

$$T_1 = -\varepsilon^{-1} (\mathbf{q} - \Pi_h^0 \mathbf{q}, \varepsilon_h^y)_{\partial T_h} \leq C \|\mathbf{q} - \Pi_h^0 \mathbf{q}\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2.$$

For the term $T_2$, the Cauchy-Schwarz inequality and an inverse inequality give

$$T_2 = (\mathbf{x}^{k+1} - \mathbf{y}, \varepsilon_h^y \cdot \mathbf{n})_{\partial T_h} \leq C h^{-\frac{1}{2}} \|\mathbf{x}^{k+1} - \mathbf{y}\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h}$$

$$\leq C h^{-\frac{1}{2}} \|\mathbf{x}^{k+1} - \mathbf{y}\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \leq C h^{-\frac{1}{2}} \|\mathbf{x}^{k+1} - \mathbf{y}\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2.$$

For the terms $T_3$, $T_4$ and $T_5$, apply (19) and Young’s inequality to obtain

$$T_3 = -(\Pi_h^0 \mathbf{q} - \mathbf{q}, \nabla \varepsilon_h^y)_{T_h} \leq C \|\Pi_h^0 \mathbf{q} - \mathbf{q}\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2,$$

$$+ C h^{-1} \|\mathbf{x}^{k+1} - \mathbf{y}\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \leq C \|\mathbf{y} - \Pi_h^k \mathbf{y}\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2,$$

$$+ C h^{-1} \|\mathbf{x}^{k+1} - \mathbf{y}\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \leq C \|\mathbf{y} - \Pi_h^k \mathbf{y}\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2.$$

For the terms $T_6$, $T_7$ and $T_8$, Young’s equality gives

$$T_6 + T_7 + T_8 \leq C h^{-1} \|\Pi_h^k \mathbf{y}\|_{\partial T_h} + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2,$$

$$+ \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2.$$

For the last two terms $T_9$ and $T_{10}$, apply the trace inequality Lemma 3.7 to get

$$T_9 + T_{10} = (\Pi_h^0 \mathbf{q} - \mathbf{q}, \varepsilon_h^y - \varepsilon_h^\gamma)_{\partial T_h} + (\Pi_h^0 \mathbf{q} - \mathbf{q}, \varepsilon_h^y)_{\partial T_h} \leq C h \|\nabla \cdot ((\Pi_h^0 \mathbf{q} - \mathbf{q}))^2_{T_h} + C h^{-1} \|\Pi_h^0 \mathbf{q} - \mathbf{q}\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2 + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2,$$

$$+ C h \|\Pi_h^0 \mathbf{q} - \mathbf{q}\|_{\partial T_h} + \frac{1}{16 \varepsilon} \|\varepsilon_h^y\|_{T_h}^2.$$
Next, we introduce the dual problem for any given $\Theta$ in $L^2(\Omega)$:
\[
\Phi + \nabla \Psi = 0 \quad \text{in } \Omega,
\]
\[
\nabla \cdot \Phi - \nabla \cdot (\beta \Psi) = \Theta \quad \text{in } \Omega,
\]
\[
\Psi = 0 \quad \text{on } \partial \Omega.
\]
(20)

Since the domain $\Omega$ is convex, we have the following regularity estimate
\[
\|\Phi\|_1 + \|\Psi\|_2 \leq C_{\text{reg}} \|\Theta\|_{\mathcal{T}_h},
\]
(21)

**Lemma 3.12** For $\varepsilon_h^\gamma$ defined in Lemma 3.10, we have
\[
\|\varepsilon_h^\gamma\|_{\mathcal{T}_h} \leq C (\|\Pi_h^{k+1} y - y\|_{\mathcal{T}_h} + h^{1/2} (\|\Pi_h^{k+1} y - y\|_{\mathcal{T}_h} + \|\Pi_h^{k+1} y - y\|_{\partial \mathcal{T}_h})))
\]

\[
+ C(h^2 \|\nabla \cdot q\|_{\mathcal{T}_h} + h \|q - \Pi_h^0 q\|_{\mathcal{T}_h}).
\]

**Proof** First, take $(r_1, w_1, \mu_1) = (\Pi_h^k \Phi, -\Pi_h^{k+1} \Psi, -\bar{T}_h^{k+1} \Psi)$ in Lemma 3.10 and use $\Psi = 0$ on $\mathcal{E}_h$ to obtain
\[
\mathcal{B}_1(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\gamma; \Pi_h^k \Phi, -\Pi_h^{k+1} \Psi, -\bar{T}_h^{k+1} \Psi)
\]
\[
= -\varepsilon_h^y (\Phi, \Pi_h^0 q, \Pi_h^k \Phi; \mathcal{T}_h) + (\bar{T}_h^{k+1} y - y, \Pi_h^k \Phi \cdot n)_{\partial \mathcal{T}_h} \varepsilon_h^\gamma
\]
\[
+ (\Pi_h^0 q - q) \cdot \Pi_h^{k+1} \Psi \Pi_h^k \Phi; \mathcal{T}_h) + (\Pi_h^k \Phi \cdot n, \nabla \Pi_h^{k+1} \Psi; \mathcal{T}_h)
\]
\[
- \langle (h^{-1} + \tau_1) (\Pi_h^{k+1} y - \bar{T}_h^{k+1} y), \Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi \rangle_{\partial \mathcal{T}_h} \varepsilon_h^\gamma
\]
\[
- \langle (h^{-1} + \tau_2) (\Pi_h^{k+1} y - y), \Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi \rangle_{\partial \mathcal{T}_h} \varepsilon_h^\gamma
\]
\[
- (\beta \cdot n (\bar{T}_h^{k+1} y - y), \Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi ; \partial \mathcal{T}_h) \varepsilon_h^\gamma
\]
\[
- (\Pi_h^0 q - q) \cdot n, \Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi ; \partial \mathcal{T}_h) \varepsilon_h^\gamma.
\]

On the other hand, Lemmas 3.5 and 3.9 imply
\[
\mathcal{B}_1(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\gamma; \Pi_h^k \Phi, -\Pi_h^{k+1} \Psi, -\bar{T}_h^{k+1} \Psi)
\]
\[
= -\mathcal{B}_2(\Pi_h^k \Phi, \Pi_h^{k+1} \Psi, -\bar{T}_h^{k+1} \Psi; -\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\gamma)
\]
\[
= -\langle (h^{-1} + \tau_2) (\Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi), \varepsilon_h^y - \varepsilon_h^\gamma \rangle_{\partial \mathcal{T}_h} \varepsilon_h^\gamma
\]
\[
- \langle (h^{-1} + \tau_2) (\Pi_h^{k+1} \Psi - \bar{T}_h^{k+1} \Psi), \varepsilon_h^y - \varepsilon_h^\gamma \rangle_{\partial \mathcal{T}_h} \varepsilon_h^\gamma.
\]
\[
+ \langle \beta \cdot n(x_h^{k+1} - y, \varepsilon_h^\gamma - \varepsilon_h^\gamma)_{\partial T_h} \rangle \\
- \langle (\Pi_h^k \Phi - \Phi) \cdot n, \varepsilon_h^\gamma - \varepsilon_h^\gamma \rangle_{\partial T_h} - \langle (\Pi_h^k \Phi - \Phi) \cdot n, \varepsilon_h^\gamma \rangle_{\partial T_h}.
\]

Compare the above two equalities and take \( \Theta = \varepsilon_h^\gamma \) to obtain
\[
\| \varepsilon_h^\gamma \|_{T_h}^2 = (\ell_h^{k+1} \Psi - \Psi, \varepsilon_h^\gamma \cdot n)_{\partial T_h} + (\beta (\Psi - \Pi_h^{k+1} \Psi), \nabla \varepsilon_h^\gamma)_{T_h} \\
- \langle (h^{-1} + \tau_2)(\Pi_h^{k+1} \Psi - X_h^{k+1} \Psi), \varepsilon_h^\gamma - \varepsilon_h^\gamma \rangle_{\partial T_h} \\
- \langle (h^{-1} + \tau_2)(\Pi_h^{k+1} \Psi - \Psi), \varepsilon_h^\gamma, \varepsilon_h^\gamma \rangle_{\partial T_h} \\
- \langle (\Pi_h^k \Phi - \Phi) \cdot n, \varepsilon_h^\gamma - \varepsilon_h^\gamma \rangle_{\partial T_h} - \langle (\Pi_h^k \Phi - \Phi) \cdot n, \varepsilon_h^\gamma \rangle_{\partial T_h} \\
+ (\beta \cdot n(x_h^{k+1} - y, \varepsilon_h^\gamma - \varepsilon_h^\gamma)_{\partial T_h} \\
+ \varepsilon^{-1}(q - \Pi^0_h q, \Pi_h^k \Phi)_{T_h} + \langle x_h^{k+1} - y, \Pi^0_h \Phi \cdot n \rangle_{\partial T_h} \\
- \langle \Pi^0_h q - q, \nabla \Pi_h^{k+1} \Psi \rangle_{T_h} - \langle \beta (y - \Pi_h^{k+1} y), \nabla \Pi_h^{k+1} \Psi \rangle_{T_h} \\
- (\nabla \cdot \beta (y - \Pi_h^{k+1} y), \Pi_h^{k+1} \Psi)_{T_h} \\
- \langle (h^{-1} + \tau_1)(\Pi_h^{k+1} y - X_h^{k+1} y), \Pi_h^{k+1} \Psi - X_h^{k+1} \Psi \rangle_{\partial T_h} \\
- \langle (h^{-1} + \tau_1)(\Pi_h^{k+1} y - y), \Pi_h^{k+1} \Psi - X_h^{k+1} \Psi \rangle_{\partial T_h} \\
- \langle \beta \cdot n(x_h^{k+1} y - y), \Pi_h^{k+1} \Psi - X_h^{k+1} \Psi \rangle_{\partial T_h} \\
- \langle (\Pi^0_h q - q) \cdot n, \Pi_h^{k+1} \Psi - X_h^{k+1} \Psi \rangle_{\partial T_h},
\]

Estimates for the above 16 terms can be easily obtained using the proof techniques in Lemma 3.11; we omit the details. We have
\[
\| \varepsilon_h^\gamma \| \leq C (\| \Pi_h^{k+1} y - y \|_{T_h} + h^{1/2}(\| \Pi_h^{k+1} y - y \|_{\partial T_h} + \| X_h^{k+1} y - y \|_{\partial T_h})) \\
+ C (h^2 \| \nabla \cdot q \|_{T_h} + h \| q - \Pi^0_h q \|_{T_h}).
\]

As a consequence, a simple application of the triangle inequality gives optimal convergence rates for \( \| q - q_h(u) \|_{T_h} \) and \( \| y - y_h(u) \|_{T_h} \):

Lemma 3.13 Let \((q, y)\) and \((q_h(u), y_h(u))\) be the solutions of (5) and (18a), respectively. We have
\[
\| q - q_h(u) \|_{T_h} \leq C \| q - \Pi^0_h q \|_{T_h} + \| \varepsilon_h^\gamma \|_{T_h} \\
\leq C \| q - \Pi^0_h q \|_{T_h} + Ch \| \nabla \cdot q \|_{T_h} \\
+ Ch^{-1/2}(\| \Pi_h^{k+1} y - y \|_{\partial T_h} + \| X_h^{k+1} y - y \|_{\partial T_h}), \tag{22a}
\]
\[
\| y - y_h(u) \|_{T_h} \leq C \| \Pi_h^{k+1} y - y \|_{T_h} + \| \varepsilon_h^\gamma \|_{T_h} \\
\leq C \| \Pi_h^{k+1} y - y \|_{T_h} + Ch^2 \| \nabla \cdot q \|_{T_h} + Ch \| q - \Pi^0_h q \|_{T_h} \\
+ Ch^{1/2}(\| \Pi_h^{k+1} y - y \|_{\partial T_h} + \| X_h^{k+1} y - y \|_{\partial T_h}). \tag{22b}
\]
3.3.4 Step 4: The Error Equation for Part 2 of the Auxiliary Problem (18b)

Subtracting part 2 of the auxiliary problem (18b) from the equality in Lemma 3.9 gives the error equation:

**Lemma 3.14** For $\varepsilon_h^p = \Pi_h^k p - p_h(u)$, $\varepsilon_h^z = I_h^{k+1} z - z_h(u)$, $\varepsilon_h^\tau = I_h^{k+1} \tau - \tau_h(u)$, we have

$$
\mathcal{B}_2(\varepsilon_h^p, \varepsilon_h^z; \varepsilon_h^\tau, r_2, w_2, \mu_2) \\
= (y - y_h(u), w_2)_{T_h} - (\Pi \cdot n, \nabla w_2)_{T_h} \\
+ (I_h^{k+1} z - z, r_2 \cdot n)_{\partial T_h \setminus \partial T_h} - (\beta(z - I_h^{k+1} z), \nabla w_2)_{T_h} \\
+ ((h^{-1} + \tau_2)(I_h^{k+1} z - I_h^{k+1} z), w_2 - \mu_2)_{\partial T_h \setminus \partial T_h} \\
+ ((I_h^{k} p - p) \cdot n, w_2 - \mu_2)_{\partial T_h \setminus \partial T_h} + ((I_h^{k} p - p) \cdot n, w_2)_{\varepsilon_h^\theta}
$$

for all $(r_2, w_2, \mu_2) \in V_h \times W_h \times M_h(o)$.

3.3.5 Step 5: Estimates for $\varepsilon_h^p$ and $\varepsilon_h^z$ by an Energy Argument

**Lemma 3.15** For $(\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^\tau)$ defined in Lemma 3.14, we have

$$
\|\varepsilon_h^p\|_T + \|\varepsilon_h^z\|_T + h^{-1} \|\varepsilon_h^\tau\|_{\partial T_h \setminus \partial T_h} + h^{-1} \|\varepsilon_h^p\|_{\varepsilon_h^\theta} \\
\leq C \|I_h^{k+1} y - y\|_T + C h^{1/2}(\|I_h^{k+1} y - y\|_{\partial T_h} + \|I_h^{k+1} y - y\|_{\partial T_h}) \\
+ C \|\nabla \cdot \eta\|_{T_h} + C \|\eta - \Pi_h^0 \eta\|_{\partial T_h} + C h^{1/2}\|p - \Pi_h^k p\|_{\partial T_h} \\
+ C\|p - \Pi_h^k p\|_{\partial T_h} + C h^{-1/2}(\|I_h^{k+1} z - z\|_{\partial T_h} + \|I_h^{k+1} z - z\|_{\partial T_h}).
$$

**Proof** First, we take $(r_2, w_2, \mu_2) = (\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^\tau)$ in Lemma 3.4 to get

$$
\mathcal{B}_2(\varepsilon_h^p, \varepsilon_h^z; \varepsilon_h^\tau, \varepsilon_h^\tau, \varepsilon_h^\tau) \\
= \varepsilon^{-1}(\varepsilon_h^p, \varepsilon_h^p, \varepsilon_h^p)_{T_h} + \left(h^{-1} + \tau_2 + \frac{1}{2} \beta \cdot n\right) (\varepsilon_h^z - \varepsilon_h^z, \varepsilon_h^z - \varepsilon_h^z)_{\partial T_h \setminus \partial T_h} \\
- \frac{1}{2} (\nabla \cdot \beta \varepsilon_h^z, \varepsilon_h^z)_{T_h} + \left(h^{-1} + \tau_2 + \frac{1}{2} \beta \cdot n\right) (\varepsilon_h^z, \varepsilon_h^z)_{\varepsilon_h^\theta}.
$$

Next, take $(r_2, w_2, \mu_2) = (\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^\tau)$ in Lemma 3.10 to obtain

$$
\mathcal{B}_2(\varepsilon_h^p, \varepsilon_h^z; \varepsilon_h^\tau, \varepsilon_h^\tau, \varepsilon_h^\tau) \\
= (\Pi_h^k p - p, \nabla \varepsilon_h^z)_{T_h} + (I_h^{k+1} z - z, \varepsilon_h^p \cdot n)_{\partial T_h} \\
- (\beta(z - I_h^{k+1} z), \nabla \varepsilon_h^z)_{T_h} + ((h^{-1} + \tau_2)(I_h^{k+1} z - I_h^{k+1} z), \varepsilon_h^z - \varepsilon_h^z)_{\partial T_h \setminus \partial T_h} \\
+ ((h^{-1} + \tau_2)(I_h^{k+1} z - z), \varepsilon_h^z)_{\varepsilon_h^\theta} - (\beta \cdot n(I_h^{k+1} z - z), \varepsilon_h^z - \varepsilon_h^z)_{\partial T_h \setminus \partial T_h} \\
+ ((\Pi_h^k p - p) \cdot n, \varepsilon_h^z - \varepsilon_h^z)_{\partial T_h \setminus \partial T_h} + ((\Pi_h^k p - p) \cdot n, \varepsilon_h^z)_{\varepsilon_h^\theta} \\
+ (y - y_h(u), \varepsilon_h^z)_{T_h} \\
= T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8.
$$
For the terms \( T_1 - T_7 \), follow the proof of Lemma 3.11 to get

\[
\sum_{i=1}^{7} T_i \leq C h^{-1}(\|z_{T_i} - z\|_{\partial T_i} + \|z - \Pi z\|_{\partial T_i}) + Ch\|\Pi P - P\|_{T_i}^2
\]

\[
+ C\|\Pi P - P\|_{T_i}^2 + \frac{1}{16h}\|\Pi P - \Pi P\|_{T_i}^2 + \frac{1}{16h}\|\Pi P - \Pi P\|_{T_i \setminus \partial T_i}^2
\]

\[
+ \frac{1}{16h}\|\Pi P - \Pi P\|_{\partial T_i}^2 + h^{-1}\|\xi^k_{T_i} + z - \Pi z\|_{\partial T_i}^2.
\]

For the last term \( T_8 \), we introduce a discrete Poincaré inequality from [4]:

\[
\|\xi^k_{h}\|_{T_i} \leq C(\|\nabla \xi^k_{h}\|_{T_i} + h^{-\frac{1}{2}}\|\xi^k_{h}\|_{T_i})
\]

\[
= C(\|\nabla \xi^k_{h}\|_{T_i} + h^{-\frac{1}{2}}\|\xi^k_{h} - \xi^k_{h}\|_{\partial T_i} + h^{-\frac{1}{2}}\|\xi^k_{h}\|_{\partial T_i})
\]

\[
\leq C(\|\nabla \xi^k_{h}\|_{T_i} + h^{-\frac{1}{2}}\|\xi^k_{h} - \xi^k_{h}\|_{\partial T_i} + h^{-\frac{1}{2}}\|\xi^k_{h}\|_{\partial T_i}).
\]  

(23)

where \(\|\xi^k_{h}\|_{\partial T_i}^2\) is the jump of \(\xi^k_{h}\) between adjacent elements and \(\|\xi^k_{h}\|_{\partial T_i} = \xi^k_{h}\) on \(\partial T_i\). Note that the above equality in (23) holds since \(\xi^k_{h}\) is single-valued on interior faces, and the last inequality in (23) holds due to the triangle inequality.

We note the inequality in (19) is valid with \((p, z, \widehat{z})\) in place of \((q, y, \widehat{y})\). This gives

\[
T_8 \leq C\|y - y\|_{T_i}^2 + \frac{1}{16}h\|\Pi P - \Pi P\|_{T_i}^2 + \frac{1}{16}h\|\Pi P - \Pi P\|_{T_i \setminus \partial T_i}^2 + \frac{1}{16}h\|\Pi P - \Pi P\|_{\partial T_i}^2.
\]

Sum the above estimates and use (23) to obtain the desired result. \(\Box\)

As a consequence, a simple application of the triangle inequality gives optimal convergence rates for \(\|p - p_h(u)\|_{T_i}\) and \(\|z - z_h(u)\|_{T_i}\):

**Lemma 3.16** Let \((p, z)\) and \((p_h(u), z_h(u))\) be the solutions of (5) and (18b), respectively. We have

\[
\|p - p_h(u)\|_{T_i} + \|z - z_h(u)\|_{T_i}
\]

\[
\leq C\|\Pi p - \Pi p\|_{T_i} + C\|\Pi z - \Pi z\|_{T_i} + C\|\Pi z - \Pi z\|_{T_i} + C\|\Pi z - \Pi z\|_{T_i}.
\]

3.3.6 Step 6: Estimate for \(\|u - u_h\|_{\partial T_i}, \|y - y_h\|_{T_i}\) and \(\|z - z_h\|_{T_i}\)

Next, we bound the error between the solutions of the auxiliary problem (18) and the discretization of the optimality system (16). This step and the next step are very similar to Steps 6 and 7 in our previous works [18,24]. We include these proofs here to make this paper self-contained.

For the remaining steps, we denote

\[
\zeta_q = q_h(u) - q_h, \ z = y_h(u) - y_h, \ \zeta^p = p_h(u) - p_h, \ z = z_h(u) - z_h
\]

(24)

(25)
Subtracting the auxiliary problem (18) and the system (16) gives the following error equations

\[ \mathcal{B}_1(\zeta_q, \zeta_y, \zeta_\gamma; r_1, w_1, \mu_1) = -(u - u_h, r_1 \cdot n - (h^{-1} + \tau_1 - \beta \cdot n)w_1)_{\mathcal{E}_h^0}, \]
\[ \mathcal{B}_2(\zeta_p, \zeta_z, \zeta_\gamma; r_2, w_2, \mu_2) = (\zeta_y, w_2)_{\mathcal{T}_h} \]

for all \((r_1, r_2, w_1, w_2, \mu_1, \mu_2) \in V_h \times V_h \times W_h \times W_h \times M_h(o) \times M_h(o)\).

**Lemma 3.17** Let \((p_h(u), z_h(u))\) be the solution of (18), \(\zeta_y\) be defined as in (24), and \(u\) and \(u_h\) be the solutions of (5) and (16), respectively. We have

\[
\gamma \| u - u_h \|^2_{E_h^0} + \| \zeta_y \|^2_{\mathcal{T}_h} = \langle \gamma u + p_h(u) \cdot n + (h^{-1} + \tau_2)z_h(u), u - u_h \rangle_{E_h^0} - \langle \gamma u_h + p_h \cdot n + (h^{-1} + \tau_2)z_h, u - u_h \rangle_{E_h^0}.
\]

**Proof** First, we have

\[
\langle \gamma u + p_h(u) \cdot n + (h^{-1} + \tau_2)z_h(u), u - u_h \rangle_{E_h^0} = \gamma \| u - u_h \|^2_{E_h^0} + \langle \zeta_p \cdot n + (h^{-1} + \tau_2)z_h, u - u_h \rangle_{E_h^0}.
\]

Next, Lemma 3.5 gives

\[
\mathcal{B}_1(\zeta_q, \zeta_y, \zeta_\gamma; \zeta_p, -\zeta_z, -\zeta_\gamma) + \mathcal{B}_2(\zeta_p, \zeta_z, \zeta_\gamma; -\zeta_q, \zeta_y, \zeta_\gamma) = 0.
\]

One the other hand, we have

\[
\mathcal{B}_1(\zeta_q, \zeta_y, \zeta_\gamma; \zeta_p, -\zeta_z, -\zeta_\gamma) + \mathcal{B}_2(\zeta_p, \zeta_z, \zeta_\gamma; -\zeta_q, \zeta_y, \zeta_\gamma) = (\zeta_y, \zeta_\gamma)_{\mathcal{T}_h} - \langle u - u_h, \zeta_p \cdot n + (h^{-1} + \tau_2)z \rangle_{E_h^0}.
\]

Comparing the above two equalities gives

\[
(\zeta_y, \zeta_\gamma)_{\mathcal{T}_h} = \langle u - u_h, \zeta_p \cdot n + (h^{-1} + \tau_2)z \rangle_{E_h^0}.
\]

\[\square\]

**Lemma 3.18** Let \((u, y)\) and \((u_h, y_h)\) be the solutions of (5) and (16), respectively. We have

\[
\| u - u_h \|^2_{E_h^0} + \| y - y_h \|_{\mathcal{T}_h} \leq Ch^{-1/2}\| \Pi_h^{k+1}y - y \|_{\mathcal{T}_h} + C\| \Pi_h^{k+1}y - y \|_{\mathcal{T}_h} + C\| \Pi_h^{k+1}y - y \|_{\mathcal{T}_h} + Ch^{-1/2}\| p - \Pi_h^{k+1}p \|_{\mathcal{T}_h} + Ch^{-1/2}\| q - \Pi_h^{k+1}q \|_{\mathcal{T}_h} + Ch^{-1/2}\| z - \Pi_h^{k+1}z \|_{\mathcal{T}_h} + C\| p - \Pi_h^{k+1}p \|_{\mathcal{T}_h} + Ch^{-1}(\| \Pi_h^{k+1}z - z \|_{\mathcal{T}_h} + \| \Pi_h^{k+1}z - z \|_{\mathcal{T}_h}).
\]

**Proof** The optimality conditions yield \(\gamma u + p \cdot n = 0\) and \(\gamma u_h + p_h \cdot n + h^{-1}z_h + \tau_2z_h = 0\) on \(E_h^0\). Therefore, the above lemma gives

\[
\gamma \| u - u_h \|^2_{E_h^0} + \| \zeta_y \|^2_{\mathcal{T}_h} = \langle \gamma u + p_h(u) \cdot n + (h^{-1} + \tau_2)z_h(u), u - u_h \rangle_{E_h^0} - \langle \gamma u_h + p_h \cdot n + (h^{-1} + \tau_2)z_h, u - u_h \rangle_{E_h^0}.
\]
Since $z = 0$ on $\mathcal{E}_h^a$, we have
\[
\| p_h(u) - p \|_{\partial T_h} \leq \| p_h(u) - \Pi_h^k p \|_{\partial T_h} + \| \Pi_h^k p - p \|_{\partial T_h} 
\leq Ch^{-\frac{1}{2}} \| e_h^p \|_{T_h} + C \| \Pi_h^k p - p \|_{\partial T_h},
\]
\[
\| z_h(u) \|_{e_h^a} = \| z_h(u) - \Pi_h^{k+1} z + \Pi_h^{k+1} z - z \|_{e_h^a} 
\leq \| e_h^z \|_{e_h^a} + \| \Pi_h^{k+1} z - z \|_{\partial T_h}.
\]
Lemma 3.15 implies
\[
\| u - u_h \|_{e_h^a} + \| \xi_y \|_{\partial T_h} 
\leq Ch^{-1/2} \| \Pi_h^{k+1} y - y \|_{T_h} + C \| \Pi_h^{k+1} y - y \|_{\partial T_h} + C \| T_h^{k+1} y - y \|_{\partial T_h} 
+ Ch^{3/2} \| \nabla \cdot q \|_{T_h} + Ch^{1/2} \| q - \Pi_h^0 q \|_{T_h} + Ch^{-1/2} \| p - \Pi_h^k p \|_{\partial T_h} 
+ Ch^{-3/2} \| z - \Pi_h^{k+1} z \|_{T_h} + C \| p - \Pi_h^k p \|_{\partial T_h} 
+ Ch^{-1} (\| \Pi_h^{k+1} z - z \|_{\partial T_h} + \| T_h^{k+1} z - z \|_{\partial T_h}).
\]
The triangle inequality and Lemma 3.13 yield the desired result. \hfill \Box

3.3.7 Step 7: Estimates for $\| p - p_h \|_{T_h}$, $\| z - z_h \|_{T_h}$, and $\| q - q_h \|_{T_h}$

**Lemma 3.19** For $(\xi_p, \xi_z)$ defined in (24), we have
\[
\| \xi_p \|_{T_h} + \| \xi_z \|_{\partial T_h} 
\leq Ch^{-1/2} \| \Pi_h^{k+1} y - y \|_{T_h} + C \| \Pi_h^{k+1} y - y \|_{\partial T_h} + C \| T_h^{k+1} y - y \|_{\partial T_h} 
+ Ch^{3/2} \| \nabla \cdot q \|_{T_h} + Ch^{1/2} \| q - \Pi_h^0 q \|_{T_h} + Ch^{-1/2} \| p - \Pi_h^k p \|_{\partial T_h} 
+ Ch^{-3/2} \| z - \Pi_h^{k+1} z \|_{T_h} + C \| p - \Pi_h^k p \|_{\partial T_h} 
+ Ch^{-1} (\| \Pi_h^{k+1} z - z \|_{\partial T_h} + \| T_h^{k+1} z - z \|_{\partial T_h}).
\]

**Proof** By the energy identity for $\mathcal{B}_2$ in Lemma 3.4, and the second error Eq. (26b), we have
\[
\mathcal{B}_2(\xi_p, \xi_z, \xi_z^a; \xi_p, \xi_z, \xi_z) 
= (\xi_y, \xi_z)_{T_h} 
\leq \| \xi_y \|_{T_h} \| \xi_z \|_{\partial T_h} 
\leq C \| \xi_y \|_{T_h} (\| \nabla \xi_z \|_{T_h} + h^{-\frac{1}{2}} \| \xi_z \|_{\partial T_h \setminus e_h^a} + h^{-\frac{1}{2}} \| \xi_z \|_{e_h^a}) 
\leq C \| \xi_y \|_{T_h} (\| \xi_p \|_{T_h} + h^{-\frac{1}{2}} \| \xi_z \|_{\partial T_h \setminus e_h^a} + h^{-\frac{1}{2}} \| \xi_z \|_{e_h^a} + h^{-\frac{1}{2}} \| \Pi_h^{k+1} z - z \|_{\partial T_h}),
\]
where for the last two inequalities we used the discrete Poincaré inequality (23) and also the inequality (19). This gives
Lemma 3.20 If $k \geq 1$, then

\[
\|\zeta_q\|_{\mathcal{T}_h} \leq C \|\Pi_h^{k+1} y - y\|_{\mathcal{T}_h} + Ch^{-1/2} \|\Pi_h^{k+1} y - y\|_{\partial \mathcal{T}_h} + Ch^{3/2} \|\nabla \cdot q\|_{\mathcal{T}_h} + Ch^{1/2} \|q - \Pi_h^0 q\|_{\mathcal{T}_h} + Ch^{-1/2} \|p - \Pi_h^k p\|_{\partial \mathcal{T}_h} + Ch^{-3/2} \|\Pi_h^{k+1} z - z\|_{\partial \mathcal{T}_h} + \|\mathcal{I}_h^{k+1} z - z\|_{\partial \mathcal{T}_h}.
\]

Using the discrete Poincaré inequality (23) and (19) again yield

\[
\|\zeta_z\|_{\mathcal{T}_h} \leq C (\|\nabla \zeta_z\|_{\mathcal{T}_h} + h^{-1/2} \|z - \zeta_z\|_{\partial \mathcal{T}_h \setminus \mathcal{E}_h} + h^{-1/2} \|z\|_{\mathcal{E}_h}).
\]

Finally, combine (27) and the above inequality to give the desired result. \hfill \Box

**Lemma 3.20** If $k = 0$ and EDG: errors, observed convergence orders, and expected order (EO) for the control $u$, the state $y$, the dual state $z$ and their fluxes $q$ and $p$

| $h/\sqrt{2}$ | 1/16 | 1/32 | 1/64 | 1/128 | 1/256 | EO |
|-----------|------|------|------|-------|-------|-----|
| $\|q - q_h\|_{\mathcal{T}_h}$ | 3.344E−01 | 2.642E−01 | 1.847E−01 | 1.191E−01 | 7.230E−02 | 2 |
| Order | – | 0.34 | 0.52 | 0.63 | 0.72 | – |
| $\|p - p_h\|_{\mathcal{T}_h}$ | 1.056E−01 | 6.199E−02 | 3.277E−02 | 1.667E−02 | 8.368E−03 | 3 |
| Order | – | 0.77 | 0.92 | 0.97 | 0.99 | 0.5 |
| $\|y - y_h\|_{\mathcal{T}_h}$ | 1.203E−01 | 6.647E-02 | 3.492E−02 | 1.768E−02 | 8.679E−03 | 3 |
| Order | – | 0.85 | 0.93 | 0.98 | 1.02 | 0.5 |
| $\|z - z_h\|_{\mathcal{T}_h}$ | 1.371E−02 | 3.955E−03 | 1.464E−03 | 6.427E−04 | 2.972E−04 | 1 |
| Order | – | 1.79 | 1.43 | 1.18 | 1.11 | 0.5 |
| $\|u - u_h\|_{\mathcal{E}_h}$ | 3.924E−01 | 2.567E−01 | 1.481E−01 | 7.930E−02 | 4.023E−02 | 1 |
| Order | – | 0.61 | 0.79 | 0.90 | 0.98 | 0.5 |

**Proof** Lemma 3.4 and the first error Eq. (26a) give

\[
\mathcal{B}_1(\zeta_q, \zeta_y, \zeta_{\tilde{y}}; \zeta_q, \zeta_y, \zeta_{\tilde{y}})
\]

\[
= e^{-1} (\zeta_q, \zeta_q)_{\mathcal{T}_h} + \left(\left(h^{-1} + \tau_1 - \frac{1}{2} \beta \cdot n\right) (\zeta_y - \zeta_{\tilde{y}}), (\zeta_y - \zeta_{\tilde{y}})_{\partial \mathcal{T}_h \setminus \mathcal{E}_h} \right)
\]

\[
- \frac{1}{2} \left(\nabla \cdot \beta \zeta_y, \zeta_y\right)_{\mathcal{T}_h} + \left(\left(h^{-1} + \tau_1 - \frac{1}{2} \beta \cdot n\right) \zeta_y, \zeta_y\right)_{\mathcal{E}_h}
\]

\[
= - (u - u_h, \zeta_q \cdot n + (\beta \cdot n - h^{-1} - \tau_1) \zeta_y)_{\mathcal{E}_h}
\]
Table 2 Example 4.1, high regularity test, $k = 0$ and IEDG: errors, observed convergence orders, and expected order (EO) for the control $u$, the state $y$, the dual state $z$ and their fluxes $q$ and $p$

| $h / \sqrt{2}$ | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 | EO |
|-----------------|-----|-----|-----|------|------|----|
| $\|q - q_h\|_{T_h}$ | 2.862E−01 | 2.051E−01 | 1.4036E−01 | 9.019E−02 | 5.552E−02 | |
| Order | – | 0.48 | 0.55 | 0.64 | 0.70 | – |
| $\|p - p_h\|_{T_h}$ | 8.072E−02 | 4.827E−02 | 2.701E−02 | 1.471E−02 | 7.754E−03 | |
| Order | – | 0.74 | 0.83 | 0.87 | 0.92 | 0.5 |
| $\|y - y_h\|_{T_h}$ | 4.608E−02 | 1.759E−02 | 5.644E−03 | 1.988E−03 | 6.866E−04 | |
| Order | – | 1.38 | 1.64 | 1.50 | 1.53 | 0.5 |
| $\|z - z_h\|_{T_h}$ | 2.053E−02 | 6.196E−03 | 1.482E−03 | 3.220E−04 | 6.930E−05 | |
| Order | – | 1.72 | 2.06 | 2.20 | 2.21 | 0.5 |
| $\|u - u_h\|_{E^d_h}$ | 1.183E−01 | 6.745E−02 | 3.904E−02 | 2.184E−02 | 1.179E−02 | |
| Order | – | 0.81 | 0.78 | 0.83 | 0.88 | 0.5 |

Table 3 Example 4.1, high regularity test, $k = 1$ and EDG: errors, observed convergence orders, and expected order (EO) for the control $u$, the state $y$, the dual state $z$ and their fluxes $q$ and $p$

| $h / \sqrt{2}$ | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 | EO |
|-----------------|-----|-----|-----|------|------|----|
| $\|q - q_h\|_{T_h}$ | 1.887E−01 | 1.056E−01 | 5.596E−02 | 2.869E−02 | 1.446E−02 | |
| Order | – | 0.83 | 0.91 | 0.96 | 0.99 | 1.0 |
| $\|p - p_h\|_{T_h}$ | 2.714E−02 | 9.111E−03 | 2.737E−03 | 7.822E−04 | 2.176E−04 | |
| Order | – | 1.57 | 1.73 | 1.80 | 1.84 | 1.5 |
| $\|y - y_h\|_{T_h}$ | 1.693E−02 | 4.892E−03 | 1.263E−03 | 3.207E−04 | 8.168E−05 | |
| Order | – | 1.79 | 1.95 | 1.97 | 1.97 | 1.5 |
| $\|z - z_h\|_{T_h}$ | 2.144E−03 | 3.460E−04 | 5.120E−05 | 7.168E−06 | 9.818E−07 | |
| Order | – | 2.63 | 2.75 | 2.83 | 2.86 | 1.5 |
| $\|u - u_h\|_{E^d_h}$ | 8.742E−02 | 3.528E−02 | 1.332E−02 | 4.856E−03 | 1.730E−03 | |
| Order | – | 1.30 | 1.40 | 1.45 | 1.48 | 1.5 |

$$= - \langle u - u_h, \xi q \cdot n - (h^{-1} + \tau_2)\xi y \rangle_{E^d_h}$$
$$= - \langle u - u_h, \xi q \cdot n - (h^{-1} + \tau_2)\xi y \rangle_{E^d_h}$$
$$\leq C \|u - u_h\|_{E^d_h} (\|\xi q\|_{E^d_h} + h^{-1} \|\xi y\|_{E^d_h})$$
$$\leq C h^{-\frac{1}{2}} \|u - u_h\|_{E^d_h} (\|\xi q\|_{T_h} + h^{-\frac{1}{2}} \|\xi y\|_{E^d_h})$$.

This gives

$$\|\xi q\|_{T_h} \leq C h^{-\frac{1}{2}} \|u - u_h\|_{E^d_h}.$$

The desired result can be obtained by the above inequality and Lemma 3.18. □

The above lemma, the triangle inequality, Lemmas 3.13 and 3.16, the estimates in (11) and Lemma 3.18 complete the proof of the main result, Theorem 3.1.
We consider three examples on a unit square domain $\Omega = [0, 1] \times [0, 1] \subset \mathbb{R}^2$, and set $\gamma = 1$ and $\beta = [-x_2^2 \sin(x_2), \cos(x_1)e^{x_2}]$. In examples 1 and 2, we computed the convergence rates without having an explicit solution of the optimality system. We numerically approximated the solution using a very fine mesh with $h = \sqrt{2} \times 2^{-9}$, and compared this reference solution against other solutions computed on meshes with larger $h$.

**Example 4.1** First, we test the high regularity case by setting $f = 0$ and $y_d = 1$. The numerical results are shown in Tables 1, 2, 3 and 4. Next, we test the low regularity case by setting $f = 0$ and $y_d = (x_1^2 + x_2^2)^{-1/3}$. The numerical results are shown in Tables 5, 6, 7 and 8.

The convergence rate for the control $u$ and the flux $q$ in Example 4.1 match our theoretical results when $k = 1$, but are higher than our theoretical results for $k = 0$. The convergence rates for other variables are higher than our theory. Similar phenomena was reported in...
Table 6 Example 4.1, low regularity test, \( k = 0 \) and IEDG: errors, observed convergence orders, and expected order (EO) for the control \( u \), the state \( y \), the dual state \( z \) and their fluxes \( q \) and \( p \)

| \( h/\sqrt{2} \) | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 | EO  |
|-------|-----|-----|-----|------|------|-----|
| \( \|q - q_h\|_{T_h} \) | 3.600E−01 | 2.788E−01 | 2.202E−01 | 1.729E−01 | 1.372E−01 |     |
| Order | −   | 0.36 | 0.34 | 0.34  | 0.33  | −   |
| \( \|p - p_h\|_{T_h} \) | 1.072E−01 | 6.488E−02 | 3.701E−02 | 2.036E−02 | 1.080E−02 |     |
| Order | −   | 0.72 | 0.80 | 0.86  | 0.91  | 0.5 |
| \( \|y - y_h\|_{T_h} \) | 5.983E−02 | 2.280E−02 | 8.000E−03 | 3.088E−03 | 1.230E−03 |     |
| Order | −   | 1.39 | 1.51 | 1.37  | 1.32  | 0.5 |
| \( \|z - z_h\|_{T_h} \) | 2.883E−02 | 9.238E−03 | 2.387E−03 | 5.672E−04 | 1.314E−04 |     |
| Order | −   | 1.64 | 1.95 | 2.07  | 2.10  | 0.5 |
| \( \|u - u_h\|_{E_h^\partial} \) | 1.511E−01 | 8.549E−02 | 5.284E−02 | 3.198E−02 | 1.911E−02 |     |
| Order | −   | 0.82 | 0.69 | 0.72  | 0.74  | 0.5 |

Table 7 Example 4.1, low regularity test, \( k = 1 \) and EDG: errors, observed convergence orders, and expected order (EO) for the control \( u \), the state \( y \), the dual state \( z \) and their fluxes \( q \) and \( p \)

| \( h/\sqrt{2} \) | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 | EO  |
|-------|-----|-----|-----|------|------|-----|
| \( \|q - q_h\|_{T_h} \) | 2.486E−01 | 1.772E−01 | 1.347E−01 | 1.041E−01 | 7.919E−02 |     |
| Order | −   | 0.48 | 0.40 | 0.37  | 0.39  | 0.33|
| \( \|p - p_h\|_{T_h} \) | 3.467E−02 | 1.305E−02 | 4.971E−03 | 1.941E−03 | 7.691E−04 |     |
| Order | −   | 1.40 | 1.39 | 1.35  | 1.33  | 0.83|
| \( \|y - y_h\|_{T_h} \) | 1.837E−02 | 5.972E−03 | 2.199E−03 | 8.997E−04 | 3.726E−04 |     |
| Order | −   | 1.62 | 1.44 | 1.28  | 1.27  | 0.83|
| \( \|z - z_h\|_{T_h} \) | 5.359E−03 | 1.333E−03 | 3.007E−04 | 6.417E−05 | 1.333E−05 |     |
| Order | −   | 2.00 | 2.14 | 2.22  | 2.26  | 0.83|
| \( \|u - u_h\|_{E_h^\partial} \) | 9.307E−02 | 4.230E−02 | 2.254E−02 | 1.288E−02 | 7.352E−03 |     |
| Order | −   | 1.13 | 0.90 | 0.80  | 0.80  | 0.83|

[18,24]. We also note that the numerically observed convergence rates are higher for IEDG for \( y \) and \( z \) in the case \( k = 0 \).

Example 4.2 Next, we demonstrate the performance of the EDG and IEDG methods in the convection dominated case. We do not compute the convergence rates here; instead for illustration we plot the state \( y_h \) in Fig. 1. Moreover, we also plot the approximate state computed using the CG method. All computations are on the same mesh with \( h = \sqrt{2} \times 2^{-8} \) and the data chosen as

\[
\varepsilon = 10^{-6}, \quad f = x_1 x_2, \quad \text{and} \quad y_d = 1.
\]

We observe that the approximate state computed by the CG method is highly oscillatory, but we only have a small oscillation near the sharp change with the EDG and IEDG methods. Furthermore, the oscillations in the IEDG solutions are slightly smaller than in the EDG solution.
Table 8  Example 4.1, low regularity test, $k = 1$ and IEDG: errors, observed convergence orders, and expected order (EO) for the control $u$, the state $y$, the dual state $z$ and their fluxes $q$ and $p$

| $h/\sqrt{2}$ | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 | EO  |
|--------------|-----|-----|-----|------|------|-----|
| $\|q - q_h\|_{T_h}$ | 2.431E−01 | 1.786E−01 | 1.372E−01 | 1.072E−01 | 8.321E−02 |     |
| Order        | –   | 0.44 | 0.38 | 0.35  | 0.36  | 0.33 |
| $\|p - p_h\|_{T_h}$ | 3.411E−02 | 1.270E−02 | 4.868E−03 | 1.905E−03 | 7.546E−04 |     |
| Order        | –   | 1.42 | 1.38 | 1.35  | 1.33  | 0.83 |
| $\|y - y_h\|_{T_h}$ | 1.535E−02 | 5.202E−03 | 1.869E−03 | 7.186E−04 | 2.848E−04 |     |
| Order        | –   | 1.56 | 1.47 | 1.37  | 1.33  | 0.83 |
| $\|z - z_h\|_{T_h}$ | 5.255E−03 | 1.277E−03 | 2.825E−04 | 5.940E−05 | 1.212E−05 |     |
| Order        | –   | 2.04 | 2.17 | 2.24  | 2.29  | 0.83 |
| $\|u - u_h\|_{E_{\partial h}}$ | 6.326E−02 | 3.179E−02 | 1.716E−02 | 9.636E−03 | 5.458E−03 |     |
| Order        | –   | 1.00 | 0.89 | 0.83  | 0.82  | 0.83 |

Fig. 1  The computed state $y_h$ by CG (left), EDG (middle) and IEDG (right)

5 Conclusion

In this work, we approximate the solution of a convection diffusion Dirichlet boundary control problem by EDG and IEDG methods. We obtained an optimal convergence rate for the control for both high regularity and low regularity cases. Instead of introducing a special projection as in [18], we used an improved trace inequality for the low regularity case. This simplified the analysis. Finally, some numerical experiments showed that the EDG and IEDG methods are suitable for convection dominated problems. It is worth mentioning that the number of degrees of freedom of EDG and IEDG methods are lower than the HDG method.

Acknowledgements  G. Chen is supported by National Natural Science Foundation of China (NSFC) grant no. 11801063, China Postdoctoral Science Foundation project nos. 2018M633339 and 2019T120828. The research Y. Zhang is partially supported by the US National Science Foundation (NSF) under Grant number DMS-1619904.

Compliance with Ethical Standards

Conflict of interest  The authors declare that they have no conflict of interest.
References

1. Apel, T., Mateos, M., Pfefferer, J., Rösch, A.: On the regularity of the solutions of Dirichlet optimal control problems in polygonal domains. SIAM J. Control Optim. 53(6), 3620–3641 (2015). https://doi.org/10.1137/140994186
2. Apel, T., Mateos, M., Pfefferer, J., Rösch, A.: Error estimates for Dirichlet control problems in polygonal domains: quasi-uniform meshes. Math. Control Relat. Fields 8(1), 217–245 (2018). https://doi.org/10.3934/mcrf.2018010
3. Benner, P., Yücel, H.: A local discontinuous Galerkin method for Dirichlet boundary control problems. Technical Report (2017). https://doi.org/10.13140/RG.2.2.16362.52164
4. Brenner, S.C.: Poincaré–Friedrichs inequalities for piecewise $H^1$ functions. SIAM J. Numer. Anal. 41(1), 306–324 (2003). https://doi.org/10.1137/S0036142902401311
5. Cai, Z., He, C., Zhang, S.: Discontinuous finite element methods for interface problems: robust a priori and a posteriori error estimates. SIAM J. Numer. Anal. 55(1), 400–418 (2017). https://doi.org/10.1137/16M1056171
6. Casas, E., Günther, A., Mateos, M.: A paradox in the approximation of Dirichlet control problems in curved domains. SIAM J. Control Optim. 49(5), 1998–2007 (2011). https://doi.org/10.1137/100794882
7. Casas, E., Raymond, J.P.: Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. SIAM J. Control Optim. 45(5), 1586–1611 (2006). https://doi.org/10.1137/050626600
8. Celiker, F., Cockburn, B., Shi, K.: A projection-based error analysis of HDG methods for Timoshenko beams. Math. Comput. 81(277), 131–151 (2012). https://doi.org/10.1090/S0025-5718-2011-02522-6
9. Chabaud, B., Cockburn, B.: Uniform-in-time superconvergence of HDG methods for the heat equation. Math. Comput. 81(277), 107–129 (2012). https://doi.org/10.1090/S0025-5718-2011-02525-1
10. Chen, G., Hu, W., Shen, J., Singler, J.R., Zhang, Y., Zheng, X.: An HDG method for distributed control of convection diffusion PDEs. J. Comput. Appl. Math. 343, 643–661 (2018). https://doi.org/10.1016/j.cam.2018.05.028
11. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. 47(2), 1319–1365 (2009). https://doi.org/10.1137/070706616
12. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.J.: Analysis of HDG methods for Stokes flow. Math. Comput. 80(274), 723–760 (2011). https://doi.org/10.1090/S0025-5718-2010-02410-X
13. Cockburn, B., Guzmán, J., Sayas, F.J.: Coupling of Raviart–Thomas and hybridizable discontinuous Galerkin methods with BEM. SIAM J. Numer. Anal. 50(5), 2778–2801 (2012). https://doi.org/10.1137/100818339
14. Cockburn, B., Guzmán, J., Soon, S.C., Stolarski, H.K.: An analysis of the embedded discontinuous Galerkin method for second-order elliptic problems. SIAM J. Numer. Anal. 47(4), 2686–2707 (2009). https://doi.org/10.1137/080726914
15. Cockburn, B., Quenneville-Bélair, V.: Uniform-in-time superconvergence of the HDG methods for the acoustic wave equation. Math. Comput. 83(285), 65–85 (2014). https://doi.org/10.1090/S0025-5718-2013-02743-3
16. Cockburn, B., Shi, K.: Conditions for superconvergence of HDG methods for Stokes flow. Math. Comput. 82(282), 651–671 (2013). https://doi.org/10.1090/S0025-5718-2012-02644-5
17. Deckelnick, K., Günther, A., Hinze, M.: Finite element approximation of Dirichlet boundary control for elliptic PDEs on two- and three-dimensional curved domains. SIAM J. Control Optim. 48(4), 2798–2819 (2009). https://doi.org/10.1137/080735369
18. Gong, W., Hu, W., Mateos, M., Singler, J., Zhang, X., Zhang, Y.: A new HDG method for Dirichlet boundary control of convection diffusion PDEs II: low regularity. SIAM J. Numer. Anal. 56(4), 2262–2287 (2018). https://doi.org/10.1137/17M1152103
19. Gong, W., Yan, N.: Mixed finite element method for Dirichlet boundary control problem governed by elliptic PDEs. SIAM J. Control Optim. 49(3), 984–1014 (2011). https://doi.org/10.1137/100795632
20. Gunzburger, M.D., Hou, L., Svobodny, T.P.: Boundary velocity control of incompressible flow with an application to viscous drag reduction. SIAM J. Control Optim. 30(1), 167–181 (1992). https://doi.org/10.1137/0330011
21. Gunzburger, M.D., Hou, L.S., Svobodny, T.P.: Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with Dirichlet controls. RAIRO Modél. Math. Anal. Numér. 25(6), 711–748 (1991). https://doi.org/10.1051/m2an/1991250607111
22. Güzey, S., Cockburn, B., Stolarski, H.K.: The embedded discontinuous Galerkin method: application to linear shell problems. Int. J. Numer. Methods Eng. 70(7), 757–790 (2007). https://doi.org/10.1002/nme.1893
23. Hou, L.S., Svobodny, T.P.: Optimization problems for the Navier–Stokes equations with regular boundary controls. J. Math. Anal. Appl. 177(2), 342–367 (1993). https://doi.org/10.1006/jmaa.1993.1262
24. Hu, W., Mateos, M., Singler, J.R., Zhang, Y.: A new HDG method for Dirichlet boundary control of convection diffusion PDEs I: high regularity. arXiv:1801.01461
25. Hu, W., Shen, J., Singler, J.R., Zhang, Y., Zheng, X.: A superconvergent hybridizable discontinuous Galerkin method for Dirichlet boundary control of elliptic PDEs (2018). arXiv:1712.02931
26. Hu, W., Shen, J., Singler, J.R., Zhang, Y., Zheng, X.: A superconvergent HDG method for distributed control of convection diffusion PDEs. J. Sci. Comput. 76(3), 1436–1457 (2018). https://doi.org/10.1007/s10915-018-0668-z
27. John, C., Wachsmuth, D.: Optimal Dirichlet boundary control of stationary Navier–Stokes equations with state constraint. Numer. Funct. Anal. Optim. 30(11–12), 1309–1338 (2009). https://doi.org/10.1080/01630560903499001
28. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in $L^2$ for a class of evolution equations. SIAM J. Control Optim. 46(5), 1726–1753 (2007). https://doi.org/10.1137/060670110
29. Mateos, M., Neitzel, I.: Dirichlet control of elliptic state constrained problems. Comput. Optim. Appl. 63(3), 825–853 (2016). https://doi.org/10.1007/s10589-015-9784-y
30. May, S., Rannacher, R., Vexler, B.: Error analysis for a finite element approximation of elliptic Dirichlet boundary control problems. SIAM J. Control Optim. 51(3), 2585–2611 (2013). https://doi.org/10.1137/080735734
31. Nguyen, N.C., Peraire, J., Cockburn, B.: A class of embedded discontinuous Galerkin methods for computational fluid dynamics. J. Comput. Phys. 302, 674–692 (2015). https://doi.org/10.1016/j.jcp.2015.09.024
32. Qiu, W., Shen, J., Shi, K.: An HDG method for linear elasticity with strong symmetric stresses. Math. Comput. 87(309), 69–93 (2018). https://doi.org/10.1090/mcom/3249
33. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. 54(190), 483–493 (1990). https://doi.org/10.2307/2008497
34. You, P., Ding, Z., Zhou, J.: Optimal boundary control of the Stokes fluids with point velocity observations. SIAM J. Control Optim. 36(3), 981–1004 (1998). https://doi.org/10.1137/S0363012996300276
35. Zhang, X., Xie, X., Zhang, S.: An optimal embedded discontinuous galerkin method for second-order elliptic problems. Comput. Methods Appl. Math. (2018). https://doi.org/10.1515/cmam-2018-0007
36. Zhang, X., Zhang, Y., Singler, J.R.: An HDG method for distributed control of convection diffusion PDEs (2018). arXiv:1712.10106
37. Zhang, X., Zhang, Y., Singler, J.R.: An HDG method for distributed control of elliptic PDEs. Adv. Appl. Math. Mech. 10(4), 797–818 (2018). arXiv:1801.02978

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.