An EDG Method for Distributed Optimal Control of Elliptic PDEs

Xiao Zhang∗ Yangwen Zhang† John R. Singler‡

Abstract

We consider a distributed optimal control problem governed by an elliptic PDE, and propose an embedded discontinuous Galerkin (EDG) method to approximate the solution. We derive optimal a priori error estimates for the state, dual state, the optimal control, and suboptimal estimates for the fluxes. We present numerical experiments to confirm our theoretical results.

1 Introduction

We consider approximating the solution of the following distributed control problem. Let \( \Omega \subset \mathbb{R}^d \) \((d \geq 2)\) be a Lipschitz polyhedral domain with boundary \( \Gamma = \partial \Omega \). The goal is to minimize

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

subject to

\[
\begin{align*}
-\Delta y &= f + u \quad \text{in } \Omega, \\
y &= g \quad \text{on } \partial \Omega,
\end{align*}
\]

It is well known that the optimal control problem \((1.1)-(1.2)\) is equivalent to the optimality system

\[
\begin{align*}
-\Delta y &= f + u \quad \text{in } \Omega, \quad \text{(1.3a)} \\
y &= g \quad \text{on } \partial \Omega, \quad \text{(1.3b)} \\
-\Delta z &= y_d - y \quad \text{in } \Omega, \quad \text{(1.3c)} \\
z &= 0 \quad \text{on } \partial \Omega, \quad \text{(1.3d)} \\
z - \gamma u &= 0 \quad \text{in } \Omega, \quad \text{(1.3e)}
\end{align*}
\]

Different numerical methods for optimal control problems governed by partial differential equations have been extensively studied by many researchers. Numerical methods that have been investigated for this kind of problem include approaches based on standard finite element methods \([1, 7, 15, 19, 28]\), mixed finite elements \([3, 5, 6, 8, 9, 21, 22]\), and discontinuous Galerkin (DG) methods \([27, 35]\).

∗College of Mathematics, Sichuan University, Chengdu, China (zhangxiaofem@163.com). X. Zhang thanks Missouri University of Science and Technology for hosting him as a visiting scholar; some of this work was completed during his research visit.

†Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO (ywzfg4@mst.edu, singlerj@mst.edu). Y. Zhang and J. Singler were supported in part by National Science Foundation grant DMS-1217122. Y. Zhang and J. Singler thank the IMA for funding research visits, during which some of this work was completed.
Recently, hybridizable discontinuous Galerkin (HDG) methods have been developed for many partial differential equations; see, e.g., [2,4,10,11,13,14,29–31,34]. HDG methods keep the advantages of DG methods and mixed methods, while also having less globally coupled unknowns. HDG methods have now also been applied to many different optimal control problems [23–25,36].

The embedded discontinuous Galerkin (EDG) methods, originally proposed in [20], are obtained from HDG methods by replacing the discontinuous finite element space for the numerical traces with a continuous space. The number of degrees of freedoms for the EDG method are much smaller than the HDG method. This gain in computational efficiency can come with a loss: for the Poisson equation, convergence rates for the EDG method are one order lower than the HDG method [12]. However, for problems with strong convection the enhanced convergence properties of HDG methods are reduced [17]. Therefore, EDG methods are competitive for such problems, and researchers have recently begun to thoroughly investigate EDG methods for various partial differential equations [16,18,26,32,33].

Our long term goal is to devise efficient and accurate methods for complicated optimal flow control problems. EDG methods have potential for such problems; therefore, as a first step, we consider an EDG method to approximate the solution of the above optimal control problem for the Poisson equation. We use an EDG method with polynomials of degree $k$ to approximate all the variables of the optimality system (1.3), i.e., the state $y$, the dual state $z$, the numerical traces, and the fluxes $q = -\nabla y$ and $p = -\nabla z$. We describe the method in Section 2, and in Section 3 we obtain the error estimates

$$\|y - y_h\|_{0,\Omega} = O(h^{k+1}),$$

$$\|y - y_h\|_{0,\Omega} = O(h^k),$$

and

$$\|u - u_h\|_{0,\Omega} = O(h^{k+1}).$$

We present numerical results in Section 4 and then briefly discuss future work.

## 2 EDG scheme for the optimal control problem

### 2.1 Notation

Throughout the paper we adopt the standard notation $W^{m,p}(\Omega)$ for Sobolev spaces on $\Omega$ with norm $\|\cdot\|_{m,p,\Omega}$ and seminorm $|\cdot|_{m,p,\Omega}$. We denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$ with norm $\|\cdot\|_{m,\Omega}$ and seminorm $|\cdot|_{m,\Omega}$, and also $H^1_0(\Omega) = \{v \in H^1(\Omega) : v = 0$ on $\partial\Omega\}$. We denote the $L^2$-inner products on $L^2(\Omega)$ and $L^2(\Gamma)$ by

$$\langle v, w \rangle = \int_{\Omega} vw \quad \forall v, w \in L^2(\Omega),$$

$$\langle v, w \rangle = \int_{\Gamma} vw \quad \forall v, w \in L^2(\Gamma).$$

Furthermore, $H(\text{div}, \Omega) = \{v \in [L^2(\Omega)]^d, \nabla \cdot v \in L^2(\Omega)\}$.

Let $T_h$ be a collection of disjoint elements that partition $\Omega$. We denote by $\partial T_h$ the set $\{\partial K : K \in T_h\}$. For an element $K$ of the collection $T_h$, 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 $K^+$ and $K^-$ of the collection $T_h$, let $e = \partial K^+ \cap \partial K^-$ denote the interior face between $K^+$ and $K^-$ if the $d-1$ Lebesgue measure of $e$
is non-zero. Let $\varepsilon_h^o$ and $\varepsilon_h^d$ denote the set of interior and boundary faces, respectively. We denote by $\varepsilon_h$ the union of $\varepsilon_h^o$ and $\varepsilon_h^d$. We finally introduce
\[
(w, v)_{T_h} = \sum_{K \in T_h} (w, v)_K, \quad \langle \zeta, \rho \rangle_{\partial T_h} = \sum_{K \in T_h} \langle \zeta, \rho \rangle_{\partial K}.
\]

Let $P^k(D)$ denote the set of polynomials of degree at most $k$ on a domain $D$. We introduce the discontinuous finite element spaces
\[
V_h := \{v \in [L^2(\Omega)]^d : v|_K \in [P^k(K)]^d, \forall K \in T_h\},
\]
\[
W_h := \{w \in L^2(\Omega) : w|_K \in P^k(K), \forall K \in T_h\},
\]
\[
M_h := \{\mu \in L^2(\varepsilon_h) : \mu|_e \in P^k(e), \forall e \in \varepsilon_h\}.
\]

Let $M_h(o)$ and $M_h(\partial)$ denote the spaces defined in the same way as $M_h$, but with $\varepsilon_h$ replaced by $\varepsilon_h^o$ and $\varepsilon_h^d$, respectively. Spatial derivatives of functions in these discontinuous finite element spaces are understood to be taken piecewise on each element $K \in T_h$.

For EDG methods, we replace the discontinuous finite element space $M_h$ for the numerical traces with the continuous finite element space $\tilde{M}_h$ defined by
\[
\tilde{M}_h := M_h \cap C^0(\varepsilon_h).
\]
The spaces $\tilde{M}_h(o)$ and $\tilde{M}_h(\partial)$ are defined in the same way as $M_h(o)$ and $M_h(\partial)$.

### 2.2 The EDG Formulation

The mixed weak form of the optimality system \[1.3a-1.3e\] is given by
\[
(q, r_1) - (y, \nabla \cdot r_1) + \langle y, r_1 \cdot n \rangle = 0, \quad (2.5a)
\]
\[
(\nabla \cdot q, w_1) = (f + u, w_1), \quad (2.5b)
\]
\[
(p, r_2) - (z, \nabla \cdot r_2) + \langle z, r_2 \cdot n \rangle = 0, \quad (2.5c)
\]
\[
(\nabla \cdot p, w_2) = (yd - y, w_2), \quad (2.5d)
\]
\[
(z - \gamma u, v) = 0, \quad (2.5e)
\]
for all $(r_1, w_1, r_2, w_2, v) \in H(\text{div}, \Omega) \times L^2(\Omega) \times H(\text{div}, \Omega) \times L^2(\Omega) \times L^2(\Omega)$. Note that the optimality condition \[2.5d\] gives $u = \gamma^{-1}z$.

The EDG method seeks approximate fluxes $q_h, p_h \in V_h$, states $y_h, z_h \in W_h$, interior element boundary traces $\tilde{q}_h^o, \tilde{z}_h^o \in \tilde{M}_h(o)$, and control $u_h \in W_h$ satisfying
\[
(q_h, r_1)_{T_h} - (y_h, \nabla \cdot r_1)_T + \langle \tilde{q}_h^o, r_1 \cdot n \rangle_{\partial T_h \setminus \varepsilon_h^o} = -\langle I_h g, r_1 \cdot n \rangle_{\varepsilon_h^o}, \quad (2.6a)
\]
\[
-(q_h, \nabla w_1)_T + \langle \tilde{q}_h^o, n \cdot w_1 \rangle_{\partial T_h} - (u_h, w_1)_T = (f, w_1)_T, \quad (2.6b)
\]
for all $(r_1, w_1) \in V_h \times W_h$, where $I_h g$ is a continuous interpolation of $g$ on $\varepsilon_h^o$,
\[
(p_h, r_2)_{T_h} - (z_h, \nabla \cdot r_2)_T + \langle \tilde{z}_h^o, r_2 \cdot n \rangle_{\partial T_h \setminus \varepsilon_h^o} = 0, \quad (2.6c)
\]
\[
-(p_h, \nabla w_2)_T + \langle \tilde{p}_h^o, n \cdot w_2 \rangle_{\partial T_h} + (y_h, w_2)_T = (yd, w_2)_T, \quad (2.6d)
\]
for all $(r_2, w_2) \in V_h \times W_h$.
\[
\langle \tilde{q}_h, n \cdot \mu_1 \rangle_{\partial T_h \setminus \varepsilon_h^o} = 0, \quad (2.6e)
\]
\[
\langle \tilde{p}_h, n \cdot \mu_2 \rangle_{\partial T_h \setminus \varepsilon_h^o} = 0, \quad (2.6f)
\]
for all \( \mu_1, \mu_2 \in \tilde{M}_h(o) \), and the optimality condition

\[
(z_h - \gamma u_h, w_3)_{\mathcal{T}_h} = 0,
\]

for all \( w_3 \in W_h \). The EDG discrete optimality condition (2.6g) gives \( u_h = \gamma^{-1}z_h \). The numerical traces on \( \partial \mathcal{T}_h \) are defined by

\[
\begin{align*}
\hat{q}_h \cdot n &= q_h \cdot n + h^{-1}(y_h - \tilde{y}_h) \quad \text{on} \quad \partial \mathcal{T}_h \setminus \varepsilon_h^0, \\
\hat{q}_h \cdot n &= q_h \cdot n + h^{-1}(y_h - I_h g) \quad \text{on} \quad \varepsilon_h^0, \\
\hat{p}_h \cdot n &= p_h \cdot n + h^{-1}(z_h - \tilde{z}_h) \quad \text{on} \quad \partial \mathcal{T}_h \setminus \varepsilon_h^0, \\
\hat{p}_h \cdot n &= p_h \cdot n + h^{-1}z_h \quad \text{on} \quad \varepsilon_h^0.
\end{align*}
\]

Our implementation of the above EDG method and the local solver is similar to the implementation of an HDG scheme for a similar problem described in detail in [23].

### 3 Error Analysis

Next, we provide a convergence analysis of the above EDG method for the optimal control problem. Throughout this section, we assume \( \Omega \) is a bounded convex polyhedral domain, the problem data satisfies \( f \in L^2(\Omega) \) and \( g \in C^0(\partial \Omega) \), \( h \leq 1 \), and the solution of the optimality system (1.3) is sufficiently smooth.

Below, we prove our main convergence result:

**Theorem 3.1.** We have

\[
\begin{align*}
\|q - q_h\|_{\mathcal{T}_h} &\lesssim h^k(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|p - p_h\|_{\mathcal{T}_h} &\lesssim h^k(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|z - z_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|u - u_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}).
\end{align*}
\]

### 3.1 Preliminary material

The convergence analysis of the EDG method for the Poisson problem without control has been performed in [12]. The authors of [12] use a special projection to split the errors are prove the convergence. We do not use the special projection from [12] in our analysis; instead, we use the standard \( L^2 \)-orthogonal projection operators \( \Pi_V \) and \( \Pi_W \) satisfying

\[
\begin{align*}
(\Pi_V q, r)_K &= (q, r)_K \quad \forall r \in \mathcal{P}_k(K), \\
(\Pi_W y, w)_K &= (y, w)_K \quad \forall w \in \mathcal{P}_k(K).
\end{align*}
\]

In the conclusion, we briefly mention future work connected to the different EDG analysis approach taken here.

We use the following well-known bounds:

\[
\begin{align*}
\|q - \Pi_V q\|_{\mathcal{T}_h} &\leq Ch^{k+1} \|q\|_{k+1, \Omega}, \\
\|y - \Pi_W y\|_{\mathcal{T}_h} &\leq Ch^{k+1} \|y\|_{k+1, \Omega}, \\
\|y - \Pi_W y\|_{\partial \mathcal{T}_h} &\leq Ch^{k+\frac{1}{2}} \|y\|_{k+1, \Omega}, \\
\|q - \Pi_V q\|_{\partial \mathcal{T}_h} &\leq Ch^{k+\frac{1}{2}} \|q\|_{k+1, \Omega}, \\
\|y - I_h y\|_{\partial \mathcal{T}_h} &\leq Ch^{k+\frac{1}{2}} \|y\|_{k+1, \Omega}, \\
\|w\|_{\partial \mathcal{T}_h} &\leq Ch^{-\frac{1}{2}} \|w\|_{\mathcal{T}_h}, \forall w \in W_h.
\end{align*}
\]
where \( I_h \) is a continuous interpolation operator, and we have the same projection error bounds for \( p \) and \( z \).

Next, define the EDG operator \( \mathcal{B} \) by

\[
\mathcal{B}(v_h, w_h; \mu_h; r_1, w_1, \mu_1) = (q_h, r_1)_{\Gamma_h} - (y_h, \nabla \cdot r_1)_{\Gamma_h} + (r_h^0, r_1 \cdot n)_{\partial \Omega_h \setminus \Gamma_h} - (q_h, \nabla w_1)_{\Gamma_h} + (q_h \cdot n + h^{-1}y_h, w_1)_{\partial \Omega_h \setminus \Gamma_h} - (h^{-1}y_h^0, w_1)_{\partial \Omega_h \setminus \Gamma_h} - (q_h \cdot n + h^{-1}(y_h - \tilde{y}_h^0), \mu_1)_{\partial \Omega_h \setminus \Gamma_h}.
\]

(3.3)

By the definition in (3.3), we can rewrite the EDG formulation of the optimality system (2.6) as follows: find \((q_h, p_h, y_h, z_h, u_h, \tilde{y}_h^0, \tilde{z}_h^0) \in V_h \times V_h \times W_h \times W_h \times \tilde{M}_h(o) \times \tilde{M}_h(o)\) such that

\[
\mathcal{B}(q_h, y_h, \tilde{y}_h^0; r_1, w_1, \mu_1) = (f + u_h, w_1)_{\Gamma_h} + (I_h g, h^{-1}w_1 - r_1 \cdot n)_{\epsilon_h^0}, \quad (3.4a)
\]

\[
\mathcal{B}(p_h, z_h, \tilde{z}_h^0; r_2, w_2, \mu_2) = (y_d - y_h, w_2)_{\Gamma_h}, \quad (3.4b)
\]

\[
(z_h - gy_h, w_3)_{\Gamma_h} = 0, \quad (3.4c)
\]

for all \((r_1, r_2, w_1, w_2, w_3, \mu_1, \mu_2) \in V_h \times V_h \times W_h \times W_h \times \tilde{M}_h(o) \times \tilde{M}_h(o)\).

Below, we present two fundamental properties of the operator \( \mathcal{B} \), and show the EDG discretization of the optimality system (3.4) has a unique solution. The strategy of the proofs of these three results is similar to our earlier HDG work [23]; we include the proofs to make this paper self-contained.

**Lemma 3.2.** For any \((v_h, w_h, \mu_h) \in V_h \times W_h \times M_h(o)\), we have

\[
\mathcal{B}(v_h, w_h, \mu_h; v_h, w_h, \mu_h) = (v_h, v_h)_{\Gamma_h} + \langle h^{-1}(w_h - \mu_h), w_h - \mu_h \rangle_{\partial \Omega_h \setminus \Gamma_h} + \langle h^{-1}w_h, w_h \rangle_{\epsilon_h^0}.
\]

**Proof.** Compute:

\[
\mathcal{B}(v_h, w_h, \mu_h; v_h, w_h, \mu_h) = (v_h, v_h)_{\Gamma_h} - (w_h, \nabla \cdot v_h)_{\Gamma_h} + \langle \mu_h, v_h \cdot n \rangle_{\partial \Omega_h \setminus \Gamma_h} - (v_h, \nabla w_h)_{\Gamma_h} + (v_h \cdot n + h^{-1}w_h, w_h)_{\partial \Omega_h \setminus \Gamma_h} - \langle h^{-1}\mu_h, w_h \rangle_{\partial \Omega_h \setminus \Gamma_h} - (v_h \cdot n + h^{-1}(w_h - \mu_h), \mu_h)_{\partial \Omega_h \setminus \Gamma_h} = (v_h, v_h)_{\Gamma_h} + \langle h^{-1}w_h, w_h \rangle_{\partial \Omega_h \setminus \Gamma_h} - \langle h^{-1}\mu_h, w_h \rangle_{\partial \Omega_h \setminus \Gamma_h} - (h^{-1}(w_h - \mu_h), \mu_h)_{\partial \Omega_h \setminus \Gamma_h} = (v_h, v_h)_{\Gamma_h} + \langle h^{-1}(w_h - \mu_h), w_h - \mu_h \rangle_{\partial \Omega_h \setminus \Gamma_h} + \langle h^{-1}w_h, w_h \rangle_{\epsilon_h^0}.\]

**Lemma 3.3.** We have

\[
\mathcal{B}(q_h, y_h, \tilde{y}_h^0; p_h, -z_h, -\tilde{z}_h^0) + \mathcal{B}(p_h, z_h, \tilde{z}_h^0; -q_h, y_h, \tilde{y}_h^0) = 0.
\]
Proof. By the definition of $\mathcal{B}$, and integration by parts:

\[
\mathcal{B}(q_h, y_h, \hat{y}_h; p_h, -z_h, -\hat{z}_h) + \mathcal{B}(p_h, z_h, \hat{z}_h; -q_h, y_h, \hat{y}_h)
= (q_h, p_h)_{\mathcal{T}_h} - (y_h, \nabla \cdot p_h)_{\mathcal{T}_h} + (\hat{y}_h, p_h \cdot n)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0 \\
+ (q_h, \nabla z_h)_{\mathcal{T}_h} - (\hat{z}_h, q_h \cdot n + h^{-1} y_h, z_h)_{\partial \mathcal{T}_h} + (h^{-1} \hat{y}_h, z_h)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0 \\
+ (q_h \cdot n + h^{-1} (y_h - \hat{y}_h), \hat{z}_h)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0 \\
- (p_h, q_h)_{\mathcal{T}_h} + (z_h, \nabla q_h)_{\mathcal{T}_h} - (\hat{z}_h, q_h \cdot n)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0 \\
- (p_h, \nabla y_h)_{\mathcal{T}_h} + (p_h \cdot n + h^{-1} z_h, y_h)_{\partial \mathcal{T}_h} - (h^{-1} \hat{z}_h, y_h)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0 \\
- (p_h \cdot n + h^{-1} (z_h - \hat{z}_h), \hat{y}_h)_{\partial \mathcal{T}_h \setminus \mathcal{E}_h}^0
= 0.
\]

$\square$

**Proposition 3.4.** There exists a unique solution of the HDG equations (3.4).

Proof. Since the system (3.4) is finite dimensional, we only need to prove solutions are unique. To do this, we show zero is the only solution of the system (3.4) for problem data $y_d = f = g = 0$.

Take $(r_1, w_1, \mu_1) = (p_h, -z_h, -\hat{z}_h)$, $(r_2, w_2, \mu_2) = (-q_h, y_h, \hat{y}_h)$, and $w_3 = z_h - \gamma u_h$ in the EDG equations (3.4a), (3.4b), and (3.4c), respectively, and sum to obtain

\[
\mathcal{B}(q_h, y_h, \hat{y}_h; p_h, -z_h, -\hat{z}_h) + \mathcal{B}(p_h, z_h, \hat{z}_h; -q_h, y_h, \hat{y}_h)
= -(y_h, y_h)_{\mathcal{T}_h} - \gamma^{-1} (z_h, z_h)_{\mathcal{T}_h}.
\]

Since $\gamma > 0$, Lemma 3.3 implies $y_h = u_h = z_h = 0$.

Next, take $(r_1, w_1, \mu_1) = (q_h, y_h, \hat{y}_h)$ and $(r_2, w_2, \mu_2) = (p_h, z_h, \hat{z}_h)$ in the EDG equations (3.4a)- (3.4b). Lemma 3.2 gives $q_h = p_h = 0$ and $\hat{y}_h = \hat{z}_h = 0$.

### 3.2 Proof of Main Result

For our proof of the convergence results, we follow the strategy in [25] and consider the EDG discretization of the optimality system with the exact optimal control fixed. This results in the following auxiliary problem: find

\[
(q_h(u), p_h(u), y_h(u), z_h(u), \hat{y}_h(u), \hat{z}_h(u)) \in V_h \times V_h \times W_h \times \tilde{M}_h(0) \times \tilde{M}_h(0)
\]

satisfying

\[
\begin{align*}
\mathcal{B}(q_h(u), y_h(u), \hat{y}_h(u); r_1, w_1, \mu_1) &= (f + u, w_1)_{\mathcal{T}_h} \\
&+ (I_{hg}, h^{-1} w_1 - r_1 \cdot n)_{\mathcal{E}_h}^0, \quad (3.5a) \\
\mathcal{B}(p_h(u), z_h(u), \hat{z}_h(u); r_2, w_2, \mu_2) &= (y_d - y_h(u), w_2)_{\mathcal{T}_h}, \quad (3.5b)
\end{align*}
\]

for all $(r_1, r_2, w_1, w_2, \mu_1, \mu_2) \in V_h \times V_h \times W_h \times \tilde{M}_h(0) \times \tilde{M}_h(0)$.

We split our proof into seven steps, and estimate the errors between the solutions of the exact optimality system, the auxiliary problem, and the EDG discretization of the optimality system.
We start with the auxiliary problem and the mixed formulation of the optimality system (2.5a)-(2.5d). In Steps 1-3 below, we estimate the errors in the state \( y \) and the flux \( q \). We split the errors with the \( L^2 \) projections and the continuous interpolation operator. We use the following notation:

\[
\begin{align*}
\delta^q &= q - \Pi_V q, & \varepsilon^q &= \Pi_V q - q_h(u), \\
\delta^y &= y - \Pi_W y, & \varepsilon^y &= \Pi_W y - y_h(u), \\
\delta^\tilde{y} &= y - I_h y, & \varepsilon^\tilde{y} &= I_h y - \tilde{y}_h(u), \\
\tilde{\delta}_1 &= \delta^q \cdot n + h^{-1}(\delta^y - \delta^\tilde{y}), & \tilde{\varepsilon}_1 &= \varepsilon^q \cdot n + h^{-1}(\varepsilon^y - \varepsilon^\tilde{y}).
\end{align*}
\]

where \( \tilde{y}_h(u) = \bar{y}_h^0(u) \) on \( \varepsilon^q \) and \( \tilde{y}_h(u) = I_h g \) on \( \varepsilon^\tilde{y} \), which implies \( \varepsilon^\tilde{y} = 0 \) on \( \varepsilon^q \).

### 3.2.1 Step 1: The error equation for part 1 of the auxiliary problem (3.5a).

**Lemma 3.5.** We have

\[
\mathcal{B}(\varepsilon^q_h, \varepsilon^y_h, \varepsilon^\tilde{y}_h; r_1, w_1, \mu_1) = -\langle \delta^\tilde{y}, r_1 \cdot n \rangle_{\partial T_h} - \langle \tilde{\delta}_1, w_1 \rangle_{\partial T_h} + \langle \tilde{\delta}_1, \mu_1 \rangle_{\partial T_h} \varepsilon^q_h. \tag{3.7}
\]

**Proof.** By the definition of the EDG operator \( \mathcal{B} \) in (3.3), we have

\[
\begin{align*}
\mathcal{B}(\Pi_V q, \Pi_W y, I_h y; r_1, w_1, \mu_1) &= (\Pi_V q, r_1)_{T_h} - (\Pi_W y, \nabla \cdot r_1)_{T_h} + (I_h y, r_1 \cdot n)_{\partial T_h} \varepsilon^q_h \\
&\quad - (\Pi_V q, \nabla w_1)_{T_h} + \langle \Pi_V q \cdot n + h^{-1}\Pi_W y, w_1 \rangle_{\partial T_h} - (h^{-1} I_h y, w_1)_{\partial T_h} \varepsilon^y_h \\
&\quad - (\Pi_V q \cdot n + h^{-1}(\Pi_W y - I_h y), \mu_1)_{\partial T_h} \varepsilon^\tilde{y}_h.
\end{align*}
\]

Using the properties of the \( L^2 \)-orthogonal projections (3.1) gives

\[
\begin{align*}
\mathcal{B}(\Pi_V q, \Pi_W y, I_h y; r_1, w_1, \mu_1) &= (q, r_1)_{T_h} - (y, \nabla \cdot r_1)_{T_h} + \langle y, r_1 \cdot n \rangle_{\partial T_h} \varepsilon^q_h - \langle \delta^\tilde{y}, r_1 \cdot n \rangle_{\partial T_h} \varepsilon^q_h \\
&\quad - (q, \nabla w_1)_{T_h} + \langle q \cdot n, w_1 \rangle_{\partial T_h} + (h^{-1} y, w_1)_{\partial T_h} \varepsilon^y_h - \langle \delta^q \cdot n + h^{-1} \delta^y, w_1 \rangle_{\partial T_h} \\
&\quad + (h^{-1} \delta^\tilde{y}, w_1)_{\partial T_h} \varepsilon^\tilde{y}_h - \langle q \cdot n, \mu_1 \rangle_{\partial T_h} \varepsilon^\tilde{y}_h + \langle \tilde{\delta}_1, \mu_1 \rangle_{\partial T_h} \varepsilon^q_h.
\end{align*}
\]

The exact solution \( q \) and \( y \) satisfies

\[
\begin{align*}
(q, r_1)_{T_h} - (y, \nabla \cdot r_1)_{T_h} + \langle y, r_1 \cdot n \rangle_{\partial T_h} &= 0, \\
-(q, \nabla w_1)_{T_h} + \langle q \cdot n, w_1 \rangle_{\partial T_h} &= (f + u, w_1)_{T_h}, \\
\langle q \cdot n, \mu_1 \rangle_{\partial T_h} &= 0,
\end{align*}
\]

and therefore

\[
\begin{align*}
\mathcal{B}(\Pi_V q, \Pi_W y, I_h y; r_1, w_1, \mu_1) &= -\langle y, r_1 \cdot n \rangle_{\partial T_h} - (\delta^\tilde{y}, r_1 \cdot n)_{\partial T_h} + (f + u, w_1)_{T_h} + (h^{-1} y, w_1)_{\partial T_h} \\
&\quad - (\delta^q \cdot n + h^{-1} \delta^y, w_1)_{\partial T_h} + (h^{-1} \delta^\tilde{y}, w_1)_{\partial T_h} + \langle \tilde{\delta}_1, \mu_1 \rangle_{\partial T_h} \varepsilon^q_h.
\end{align*}
\]

Subtracting equation (3.5a) from the above equation completes the proof. \( \square \)
3.2.2 Step 2: Estimate for $\varepsilon_h^q$.  
**Lemma 3.6.** We have
\[
\|\varepsilon_h^q\|_{T_h} + h^{-\frac{1}{2}}\|\varepsilon_h^y - \varepsilon_h^\hat{y}\|_{\partial T_h} \lesssim h^k(|q|_{k+1} + |y|_{k+1}).
\] (3.8)

**Proof.** Take $(r_1, w_1, \mu_1) = (\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\hat{y})$ in equation (3.7) and use $\varepsilon_h^\hat{y} = 0$ on $\varepsilon_h$ to get
\begin{align*}
\mathcal{B}(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\hat{y}) &= -\langle \delta \varepsilon_h^q, \varepsilon_h^y \cdot n \rangle_{\partial T_h} - \langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h} \\
&= -\langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h}.
\end{align*}

Next, we have
\begin{align*}
-\langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h} &\leq C\|\delta \varepsilon_h^y\|_{\partial T_h} \|\varepsilon_h^q\|_{\partial T_h} \leq C h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} \|\varepsilon_h^q\|_{T_h}, \\
\langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h} &= -\langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h} - \langle \delta \varepsilon_h^y, \varepsilon_h^\hat{y} \rangle_{\partial T_h} \\
&\leq (\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{-1}\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{-1}\|\delta \varepsilon_h^y\|_{\partial T_h})\|\varepsilon_h^y - \varepsilon_h^\hat{y}\|_{\partial T_h} \\
&= (h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h})h^{-\frac{1}{2}}\|\varepsilon_h^y - \varepsilon_h^\hat{y}\|_{\partial T_h}.
\end{align*}

The energy property of operator $\mathcal{B}$ in Lemma 3.2 gives
\[
\|\varepsilon_h^q\|_{T_h} + h^{-\frac{1}{2}}\|\varepsilon_h^y - \varepsilon_h^\hat{y}\|_{\partial T_h} \lesssim h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{-\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} + h^{\frac{1}{2}}\|\delta \varepsilon_h^y\|_{\partial T_h} \\
\lesssim h^k(|q|_{k+1} + |y|_{k+1}).
\]
\[ \square \]

3.2.3 Step 3: Estimate for $\varepsilon_h^y$ by a duality argument.

Next, we introduce the dual problem for any given $\Theta$ in $L^2(\Omega)$:
\[
\begin{align*}
\Phi + \nabla \Psi &= 0 & \text{in } \Omega, \\
\nabla \cdot \Phi &= \Theta & \text{in } \Omega, \\
\Psi &= 0 & \text{on } \partial \Omega.
\end{align*}
\] (3.9)

Since the domain $\Omega$ is convex, we have the regularity estimate
\[
\|\Phi\|_{1,\Omega} + \|\Psi\|_{2,\Omega} \leq C_{\text{reg}} \|\Theta\|_{\Omega}.
\] (3.10)

In the proof below for estimating $\varepsilon_h^y$, we use the following notation:
\[
\delta \Phi = \Phi - \Pi_V \Phi, \quad \delta \Psi = \Psi - \Pi_W \Psi, \quad \delta \hat{\Psi} = \hat{\Psi} - I_h \hat{\Psi}.
\] (3.11)

**Lemma 3.7.** We have
\[
\|\varepsilon_h^y\|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}).
\] (3.12)

**Proof.** First, take $(r_1, w_1, \mu_1) = (\Pi_V \Phi, -\Pi_W \Psi, -I_h \Psi)$ in equation (3.7) to get
\begin{align*}
\mathcal{B}(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^\hat{y}) = (\varepsilon_h^q, \Pi_V \Phi)_{T_h} - (\varepsilon_h^q, \nabla \cdot \Pi_V \Phi)_{T_h} + (\varepsilon_h^\hat{y}, \Pi_V \Phi)_{T_h \\ n} \\
&+ (\varepsilon_h^q, \nabla \Pi_W \Psi)_{T_h} - (\varepsilon_h^q, \nabla \cdot \Pi_W \Psi)_{T_h} + (\varepsilon_h^\hat{y}, \Pi_W \Psi)_{T_h \\ n} \\
&+ (\varepsilon_h^y, \nabla \cdot \Pi_W \Psi)_{T_h} - (\varepsilon_h^y, \Pi_W \Psi)_{T_h \\ n} \\
&+ (\varepsilon_h^\hat{y}, -I_h \Psi)_{T_h \\ \mu}.
\end{align*}
Next, integration by parts gives
\[
-(\varepsilon_h^y, \nabla \cdot \Pi V \Phi)_{\partial T_h} = (\nabla \varepsilon_h^y, \Phi)_{\partial T_h} - (\varepsilon_h^y, \Pi V \Phi \cdot n)_{\partial T_h}
\]
\[
= -(\varepsilon_h^y, \nabla \cdot \Phi)_{\partial T_h} + (\varepsilon_h^y, \delta \Phi \cdot n)_{\partial T_h},
\]
\[
(\varepsilon_h^q, \nabla \Pi W \Psi)_{\partial T_h} = -(\nabla \cdot \varepsilon_h^q, \Psi)_{\partial T_h} + (\varepsilon_h^q \cdot n, \Pi W \Psi)_{\partial T_h}
\]
\[
= (\varepsilon_h^q, \nabla \Psi)_{\partial T_h} - (\varepsilon_h^q \cdot n, \delta \Psi)_{\partial T_h}.
\]
Since $\Phi$ and $\Psi$ satisfy the dual problem (3.9) with $\Theta = -\varepsilon_h^y$, we obtain
\[
B(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^y; \Pi V \Phi, -\Pi W \Psi, -I_h \Psi)
\]
\[
= (\varepsilon_h^q, \Phi)_{\partial T_h} - (\varepsilon_h^y, \nabla \Phi)_{\partial T_h} + (\varepsilon_h^y, \delta \Phi \cdot n)_{\partial T_h}
\]
\[
+ (\varepsilon_h^q, \nabla \Psi)_{\partial T_h} - (\varepsilon_h^q \cdot n, \Psi)_{\partial T_h} - \langle h^{-1} \varepsilon_h^y, \Pi W \Psi \rangle_{\partial T_h}
\]
\[
+ \langle h^{-1} \varepsilon_h^y, \Pi W \Psi \rangle_{\partial T_h} + (\varepsilon_h^q \cdot n + h^{-1}(\varepsilon_h^y - \varepsilon_h^y), I_h \Psi)_{\partial T_h}
\]
\[
= (\varepsilon_h^y, \varepsilon_h^y)_{\partial T_h} + (\varepsilon_h^y - \varepsilon_h^y, \delta \Phi \cdot n)_{\partial T_h} - (\varepsilon_h^q \cdot n, \delta \Psi)_{\partial T_h}
\]
\[
+ \langle h^{-1} (\varepsilon_h^y - \varepsilon_h^y), \delta \Psi - \delta \Psi \rangle_{\partial T_h},
\]
where we used $\langle \varepsilon_h^y, \Phi \cdot n \rangle_{\partial T_h \setminus \varepsilon_h^y} = 0$ and $\Psi = \delta \Psi = 0$ on $\varepsilon_h^y$.

On the other hand, from equation (3.7) and $\langle \delta \Psi, \Phi \cdot n \rangle_{\partial T_h} = 0$ we have
\[
B(\varepsilon_h^q, \varepsilon_h^y, \varepsilon_h^y; \Pi V \Phi, -\Pi W \Psi, -I_h \Psi)
\]
\[
= -\langle \delta \Psi, \Pi V \Phi \cdot n \rangle_{\partial T_h} + (\delta_1, \delta \Psi - \delta \Psi)_{\partial T_h}
\]
\[
= \langle \delta \Psi, \delta \Phi \cdot n \rangle_{\partial T_h} + (\delta_1, \delta \Psi - \delta \Psi)_{\partial T_h}.
\]
Comparing with the two equations above, we have
\[
\|\varepsilon_h^y\|^2_{T_h} = \langle \delta \Psi, \delta \Phi \cdot n \rangle_{\partial T_h} + (\delta_1, \delta \Psi - \delta \Psi)_{\partial T_h} - \langle \varepsilon_h^y - \varepsilon_h^y, \delta \Phi \cdot n \rangle_{\partial T_h}
\]
\[
+ \langle \varepsilon_h^q \cdot n, \delta \Psi \rangle_{\partial T_h} - \langle h^{-1} (\varepsilon_h^y - \varepsilon_h^y), \delta \Psi - \delta \Psi \rangle_{\partial T_h}
\]
\[
=: T_1 + T_2 + T_3 + T_4 + T_5.
\]
By the Cauchy-Schwarz inequality, Lemma 3.6 and (3.3), we have
\[
T_1 \leq \|\delta \Psi\|_{\partial T_h} \|\delta \Phi\|_{\partial T_h} \lesssim h^{\frac{1}{2}} \|\delta \Psi\|_{\partial T_h} \|\Phi\|_{1,\Omega} \lesssim h^k \|\delta \Psi\|_{\partial T_h} \|\varepsilon_h^y\|_{\Omega},
\]
\[
\lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}) \|\varepsilon_h^y\|_{T_h},
\]
\[
T_2 \lesssim h^{\frac{1}{2}} \|\delta \Psi\|_{\partial T_h} + h^{-1} \|\delta \Psi - \delta \Psi\|_{\partial T_h} \|\psi\|_{2,\Omega}
\]
\[
\lesssim \langle h \|\delta \Psi\|_{\partial T_h} + h^{\frac{1}{2}} \|\delta \Psi\|_{\partial T_h} \|\delta \Psi - \delta \Psi\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h}
\]
\[
\lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}) \|\varepsilon_h^y\|_{T_h},
\]
\[
T_3 \lesssim \langle \varepsilon_h^y - \varepsilon_h^y, \delta \Phi \cdot n \rangle_{\partial T_h} \lesssim h^{\frac{1}{2}} \|\varepsilon_h^y - \varepsilon_h^y\|_{\partial T_h} \|\Phi\|_{1,\Omega}
\]
\[
\lesssim h^{1/2} \|\varepsilon_h^y - \varepsilon_h^y\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}) \|\varepsilon_h^y\|_{T_h},
\]
\[
T_4 \lesssim \langle h \|\varepsilon_h^y\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}) \|\varepsilon_h^y\|_{T_h},
\]
\[
T_5 \lesssim h^{\frac{1}{2}} \|\varepsilon_h^y - \varepsilon_h^y\|_{\partial T_h} \|\varepsilon_h^y\|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}) \|\varepsilon_h^y\|_{T_h}.
\]
Summing $T_1$ to $T_5$ gives
\[ \| \varepsilon_h^y \|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}). \]

The triangle inequality gives convergence rates for $\| q - q_h(u) \|_{T_h}$ and $\| y - y_h(u) \|_{T_h}$:

**Lemma 3.8.**
\[ \| q - q_h(u) \|_{T_h} \leq \| \delta q \|_{T_h} + \| \varepsilon_h^q \|_{T_h} \lesssim h^k(|q|_{k+1} + |y|_{k+1}), \tag{3.13a} \]
\[ \| y - y_h(u) \|_{T_h} \leq \| \delta p \|_{T_h} + \| \varepsilon_h^p \|_{T_h} \lesssim h^{k+1}(|q|_{k+1} + |y|_{k+1}). \tag{3.13b} \]

### 3.2.4 Step 4: The error equation for part 2 of the auxiliary problem (3.5b)

Next, we consider the dual equation (2.5c)-(2.5d) in the optimality system and compare with the second part of the auxiliary EDG equation (3.5b). We split the errors as before; define
\[
\begin{align*}
\delta^p &= p - \Pi_V p, & \varepsilon_h^p &= \Pi_V p - p_h(u), \\
\delta^z &= z - \Pi_W z, & \varepsilon_h^z &= \Pi_W z - z_h(u), \\
\delta \hat{z} &= z - I_h z, & \varepsilon_h^{\hat{z}} &= I_h z - \hat{z}_h(u), \\
\tilde{\delta}_2 &= \delta^p \cdot n + h^{-1}(\delta^z - \delta \hat{z}), & \tilde{\varepsilon}_2 &= \varepsilon_h^p \cdot n + h^{-1}(\varepsilon_h^z - \varepsilon_h^{\hat{z}}),
\end{align*}
\]
where $\hat{z}_h(u) = \hat{z}_h^0(u)$ on $\varepsilon_h^0$ and $\hat{z}_h(u) = 0$ on $\varepsilon_h^\theta$. This gives $\varepsilon_h^{\hat{z}} = 0$ on $\varepsilon_h^\theta$.

**Lemma 3.9.** We have
\[ \mathcal{B}(\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^{\hat{z}}; r_2, w_2, \mu_2) = \langle \delta \hat{z}, r_2 \cdot n \rangle_{\partial T_h} + (y_h(u) - y, w_2)_{T_h} + (\tilde{\delta}_2, w_2 - \mu_2)_{\partial T_h}. \tag{3.15} \]

The proof is similar to the proof of Lemma 3.9 and is omitted.

### 3.2.5 Step 5: Estimates for $\varepsilon_h^p$ and $\varepsilon_h^z$ by an energy and duality argument.

**Lemma 3.10.** Let $\kappa$ be any positive constant. Then there exists a constant $C$ that does not depend on $\kappa$ such that
\[ \| \varepsilon_h^p \|_{T_h} + h^{-\frac{1}{2}}\| \varepsilon_h^z - \varepsilon_h^{\hat{z}} \|_{\partial T_h} \leq E + \kappa \| \varepsilon_h^\theta \|_{T_h}, \tag{3.16} \]
where
\[ E = Ch^{-\frac{1}{2}}\| \delta \hat{z} \|_{\partial T_h} + Ch^{-\frac{1}{2}}\| \delta^z \|_{\partial T_h} + \frac{C}{\kappa} |y_h(u) - y|_{T_h} + C\| \delta^p \|_{T_h}. \]

**Proof.** Taking $(r_2, w_2, \mu_2) = (\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^{\hat{z}})$ in (3.15) in Lemma 3.9 gives
\[
\mathcal{B}(\varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^{\hat{z}}; \varepsilon_h^p, \varepsilon_h^z, \varepsilon_h^{\hat{z}}) = \langle \delta \hat{z}, \varepsilon_h^p \cdot n \rangle_{\partial T_h} + (y_h(u) - y, \varepsilon_h^z)_{T_h} + (\tilde{\delta}_2, \varepsilon_h^{\hat{z}} - \varepsilon_h^{\hat{z}})_{\partial T_h}
\]
\[
\leq Ch^{-\frac{1}{2}}\| \delta \hat{z} \|_{\partial T_h} \| \varepsilon_h^p \|_{T_h} + \frac{1}{\kappa} |y_h(u) - y|_{T_h}^2 + \kappa \| \varepsilon_h^\theta \|_{T_h}^2
\]
\[
+ C(h^\frac{1}{2} \| \delta^p \|_{T_h} + h^{-\frac{1}{2}}\| \delta^z - \delta \hat{z} \|_{\partial T_h} + h^{-\frac{1}{2}}\| \varepsilon_h^z - \varepsilon_h^{\hat{z}} \|_{\partial T_h}).
\]
Lemma 3.2 gives
\[
\|\varepsilon_h^P\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|\varepsilon_h^\circ - \varepsilon_h^\circ\|_{\mathcal{O}_{\mathcal{T}_h}} \\
\leq Ch^{-\frac{1}{2}}\|\delta\varepsilon\|_{\mathcal{O}_{\mathcal{T}_h}} + Ch^{-\frac{1}{2}}\|\delta\varepsilon\|_{\mathcal{O}_{\mathcal{T}_h}} + \frac{C}{\kappa}\|y_h(u) - y\|_{\mathcal{T}_h} \\
+ C\|\delta P\|_{\mathcal{T}_h} + \kappa\|\varepsilon_h^\circ\|_{\mathcal{T}_h},
\]
where \(\kappa\) is any positive constant. \(\square\)

Lemma 3.11. We have
\[
\|\varepsilon_h^P\|_{\mathcal{T}_h} \lesssim h^k(||q||_{k+1} + ||y||_{k+1} + ||p||_{k+1} + ||z||_{k+1}) \tag{3.17a}
\]
\[
\|\varepsilon_h^\circ\|_{\mathcal{T}_h} \lesssim h^{k+1}(||q||_{k+1} + ||y||_{k+1} + ||p||_{k+1} + ||z||_{k+1}) \tag{3.17b}
\]

Proof. First, take \((r_2, w_2, \mu_2) = (\Pi V \Phi, -\Pi W \Psi, -I_h \psi)\) in \([3.15]\) in Lemma 3.9 to obtain
\[
\mathcal{B}(\varepsilon_h^P, \varepsilon_h^\circ, \varepsilon_h^\circ; \Pi V \Phi, -\Pi W \Psi, -I_h \psi) \\
= \langle \delta\varepsilon, \Pi V \Phi \cdot n \rangle_{\mathcal{O}_{\mathcal{T}_h}} - (y_h(u) - y, \Pi W \Psi)_{\mathcal{T}_h} - \langle \delta_2, \delta\psi - \delta\varepsilon \rangle_{\mathcal{O}_{\mathcal{T}_h}}.
\]
Next, consider the dual problem \([3.9]\) and let \(\Theta = -\varepsilon_h^\circ\). Using the definition of \(\mathcal{B}\) and the proof technique for Lemma 3.7 gives
\[
\mathcal{B}(\varepsilon_h^P, \varepsilon_h^\circ, \varepsilon_h^\circ; \Pi V \Phi, -\Pi W \Psi, -I_h \psi) \\
= (\varepsilon_h^P, \varepsilon_h^\circ)_{\mathcal{T}_h} + (\varepsilon_h^\circ - \varepsilon_h^\circ, \delta\Phi \cdot n)_{\mathcal{O}_{\mathcal{T}_h}} - (\varepsilon_h^P, \delta\psi)_{\mathcal{O}_{\mathcal{T}_h}} \\
+ (h^{-1}(\varepsilon_h^\circ - \varepsilon_h^\circ), \delta\psi - \delta\varepsilon)_{\mathcal{O}_{\mathcal{T}_h}}.
\]
Here, we used \(\langle \varepsilon_h^\circ, \Phi \cdot n \rangle_{\mathcal{O}_{\mathcal{T}_h}} = 0\), which holds since \(\varepsilon_h^\circ\) is a single-valued function on interior edges and \(\varepsilon_h^\circ = 0\) on \(\varepsilon_h^\circ\).
Comparing the above two equalities gives
\[
\|\varepsilon_h^\circ\|_{\mathcal{T}_h}^2 = (\delta\varepsilon, \Pi V \Phi \cdot n)_{\mathcal{O}_{\mathcal{T}_h}} - (y_h(u) - y, \Pi W \Psi)_{\mathcal{T}_h} - \langle \delta_2, \delta\psi - \delta\varepsilon \rangle_{\mathcal{O}_{\mathcal{T}_h}} \\
- (\varepsilon_h^\circ - \varepsilon_h^\circ, \delta\Phi \cdot n)_{\mathcal{O}_{\mathcal{T}_h}} + (\varepsilon_h^P, \delta\psi)_{\mathcal{O}_{\mathcal{T}_h}} \\
- (h^{-1}(\varepsilon_h^\circ - \varepsilon_h^\circ), \delta\psi - \delta\varepsilon)_{\mathcal{O}_{\mathcal{T}_h}},
\]
\[
= \sum_{i=1}^6 R_i.
\]
Let \(C_0 = \max\{C, 1\}\), where \(C\) is the constant defined in \([3.2]\). For the terms \(R_1, R_2,\) and \(R_3\), we have
\[
R_1 = -\langle \delta\varepsilon, \delta\Phi \cdot n \rangle_{\mathcal{O}_{\mathcal{T}_h}} \leq C_0 h^\frac{3}{2} \|\delta\varepsilon\|_{\mathcal{O}_{\mathcal{T}_h}} \|\Phi\|_{1, \Omega} \leq C_0 C_{\text{reg}} h^\frac{3}{2} \|\delta\varepsilon\|_{\mathcal{O}_{\mathcal{T}_h}} \|\varepsilon_h^\circ\|_{\mathcal{T}_h},
\]
\[
R_2 \leq \|y_h(u) - y\|_{\mathcal{T}_h} \|\delta\Phi\|_{\mathcal{T}_h} + \|\Psi\|_{\Omega} \leq C_0 C_{\text{reg}} \|y - y_h(u)\|_{\mathcal{T}_h} \|\varepsilon_h^\circ\|_{\mathcal{T}_h},
\]
\[
R_3 \leq C_0 h^\frac{3}{2} (\delta\Phi \cdot n + \frac{1}{h} \|\delta\varepsilon - \delta\varepsilon \|_{\mathcal{O}_{\mathcal{T}_h}}) \|\Psi\|_{2, \Omega} \\
\leq C_0 C_{\text{reg}} (h \|\delta\Phi\|_{\mathcal{T}_h} + h^\frac{3}{2} \|\delta\varepsilon - \delta\varepsilon \|_{\mathcal{O}_{\mathcal{T}_h}}) \|\varepsilon_h^\circ\|_{\mathcal{T}_h}.
\]
For the terms $R_4$, $R_5$ and $R_6$, Lemma 3.10 gives

$$R_4 = C_0 h^{\frac{1}{2}} \| \varepsilon^x_h - \varepsilon^x_h \|_{\partial \Omega_h} \| \Phi \|_{1, \Omega} \leq C_0 C_{\text{reg}} h (E + \kappa \| \varepsilon^x_h \|_{\Omega_h}) \| \varepsilon^x_h \|_{\Omega_h},$$

$$R_5 \leq C_0 h^{\frac{1}{2}} \| \varepsilon^p_h \|_{\partial \Omega_h} \| \Psi \|_{2, \Omega} \leq C_0 C_{\text{reg}} h \| \varepsilon^p_h \|_{\Omega_h} \| \varepsilon^x_h \|_{\Omega_h},$$

$$R_6 \leq C_0 h^{\frac{1}{2}} \| \varepsilon^x_h - \varepsilon^x_h \|_{\partial \Omega_h} \| \Psi \|_{2, \Omega} \leq C_0 C_{\text{reg}} h^{\frac{1}{2}} (E + \kappa \| \varepsilon^x_h \|_{\Omega_h}) \| \varepsilon^x_h \|_{\Omega_h}.$$ 

Summing $R_1$ to $R_6$ gives

$$\| \varepsilon^x_h \|_{\Omega_h} \leq C (h \| \delta^p \|_{\Omega_h} + \| y - y_h(u) \|_{\Omega_h} + h^{\frac{1}{2}} \| \delta^z \|_{\partial \Omega_h} + h^{\frac{1}{2}} \| \delta^z \|_{\partial \Omega_h})$$

$$+ C (E + \kappa \| \varepsilon^x_h \|_{\Omega_h}),$$

where $C = 3C_0 C_{\text{reg}}$. Choose $\kappa = \frac{1}{2C}$ gives

$$\| \varepsilon^x_h \|_{\Omega_h} \lesssim h^k \left( |q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1} \right).$$

Finally, (3.16) and (3.17b) imply (3.17a). \hfill \Box

The triangle inequality again gives convergence rates for $\| p - p_h(u) \|_{\Omega_h}$ and $\| z - z_h(u) \|_{\Omega_h}$:

**Lemma 3.12.**

$$\| p - p_h(u) \|_{\Omega_h} \leq \| \delta^p \|_{\Omega_h} + \| \varepsilon^p_h \|_{\Omega_h}$$

$$\lesssim h^k \left( |q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1} \right), \quad (3.18a)$$

$$\| z - z_h(u) \|_{\Omega_h} \leq \| \delta^z \|_{\Omega_h} + \| \varepsilon^z_h \|_{\Omega_h}$$

$$\lesssim h^k \left( |q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1} \right). \quad (3.18b)$$

### 3.2.6 Step 6: Estimates for $\| u - u_h \|_{\Omega_h}$, $\| y - y_h \|_{\Omega_h}$, and $\| z - z_h \|_{\Omega_h}$.

Next, we compare the auxiliary problem to the EDG discretization of the optimality system [3.4]. The resulting error bounds along with the earlier error bounds in Lemma 3.8 and Lemma 3.12 give the main convergence result.

The proofs in the final steps are similar to the HDG work [23]; we include the proofs here for completeness.

For the remainder of the proof, let

$$\zeta_q = q_h(u) - q_h, \quad \zeta_y = y_h(u) - y_h, \quad \zeta_{\hat{y}} = \hat{y}_h(u) - \hat{y}_h,$$

$$\zeta_p = p_h(u) - p_h, \quad \zeta_z = z_h(u) - z_h, \quad \zeta_{\hat{z}} = \hat{z}_h(u) - \hat{z}_h.$$

Subtracting the auxiliary problem and the EDG problem yields the error equations

$$B(\zeta_q, \zeta_y, \zeta_{\hat{y}}; r_1, w_1, \mu_1) = (u - u_h, w_1)_{\Omega_h}, \quad (3.19a)$$

$$B(\zeta_p, \zeta_z, \zeta_{\hat{z}}; r_2, w_2, \mu_2) = - (\zeta_y, w_2)_{\Omega_h}. \quad (3.19b)$$

**Lemma 3.13.** We have

$$\gamma \| u - u_h \|_{\Omega_h}^2 + \| y_h(u) - y_h \|_{\Omega_h}^2$$

$$= (z_h - \gamma u_h, u - u_h)_{\Omega_h} - (z_h(u) - \gamma u, u - u_h)_{\Omega_h}. \quad (3.20)$$

12
Proof. First,

\[(z_h - \gamma u_h, u - u_h)_{\mathcal{T}_h} - (z_h(u) - \gamma u, u - u_h)_{\mathcal{T}_h} = -\langle \zeta, u - u_h \rangle_{\mathcal{T}_h} + \gamma \|u - u_h\|^2_{\mathcal{T}_h} \]

Next, Lemma 3.3 and (3.19) give

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

This gives \(-\langle u - u_h, \zeta \rangle_{\mathcal{T}_h} - \|\zeta_y\|^2_{\mathcal{T}_h}\), which completes the proof.

\[\square\]

**Theorem 3.14.** We have

\[
\begin{align*}
\|u - u_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|z - z_h\|_{\mathcal{T}_h} &\lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1})
\end{align*}
\]

Proof. As mentioned earlier, the exact and approximate optimal controls satisfy \(\gamma u = z\) and \(\gamma u_h = z_h\); see (1.3c) and (3.4c). Using these equations with the lemma above give

\[
\begin{align*}
\gamma \|u - u_h\|^2_{\mathcal{T}_h} + \|\zeta_y\|^2_{\mathcal{T}_h} &= (z_h - \gamma u_h, u - u_h)_{\mathcal{T}_h} - (z_h(u) - \gamma u, u - u_h)_{\mathcal{T}_h} \\
&= -\langle z_h(u) - z, u - u_h \rangle_{\mathcal{T}_h} \\
&\leq \|z_h(u) - z\|_{\mathcal{T}_h} \|u - u_h\|_{\mathcal{T}_h} \\
&\leq \frac{1}{2\gamma} \|z_h(u) - z\|^2_{\mathcal{T}_h} + \gamma \|u - u_h\|^2_{\mathcal{T}_h}
\end{align*}
\]

Lemma 3.12 gives

\[
\|u - u_h\|_{\mathcal{T}_h} + \|\zeta_y\|_{\mathcal{T}_h} \lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}).
\]

Use the triangle inequality and Lemma 3.8 to obtain

\[
\|y - y_h\|_{\mathcal{T}_h} \lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}).
\]

Finally, the above estimate (3.22) for \(u\) along with \(z = \gamma u\) and \(z_h = \gamma u_h\) give the estimate (3.21c) for \(z\).

\[\square\]

### 3.2.7 Step 7: Estimates for \(\|q - q_h\|_{\mathcal{T}_h}\) and \(\|p - p_h\|_{\mathcal{T}_h}\).

**Lemma 3.15.** We have

\[
\begin{align*}
\|\zeta_q\|_{\mathcal{T}_h} &\lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
\|\zeta_p\|_{\mathcal{T}_h} &\lesssim h^{k+1}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1})
\end{align*}
\]

Proof. Lemma 3.2 the error equation (3.19a), and the estimate (3.22) give

\[
\begin{align*}
\|\zeta_q\|_{\mathcal{T}_h}^2 &\lesssim \mathcal{B}(\zeta_q, \zeta_y, \zeta_y; \zeta_q, \zeta_y, \zeta_y) \\
&= (u - u_h, \zeta_q)_{\mathcal{T}_h} \\
&\leq \|u - u_h\|_{\mathcal{T}_h} \|\zeta_q\|_{\mathcal{T}_h} \\
&\lesssim h^{2k+2}(\|q\|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1})^2.
\end{align*}
\]
Similarly, Lemma 3.2, the error equation (3.19b), the estimate (3.22), Lemma 3.12, and Theorem 3.14 give

\[
\| \zeta_p \|_{T_h}^2 \lesssim \mathcal{B}(\zeta_p, \zeta_z; \zeta_p, \zeta_z, \zeta_{\tilde{z}}) \\
= - (\zeta_{\hat{y}}, \zeta_{\hat{z}})_{T_h} \\
\leq \| \zeta_y \|_{T_h} \| \zeta_z \|_{T_h} \\
\leq \| \zeta_y \|_{T_h} (\| z_h(u) - z \|_{T_h} + \| z - z_h \|_{T_h}) \\
\lesssim h^{2k+2}(\| q \|_{k+1} + | y |_{k+1} + | p |_{k+1} + | z |_{k+1})^2.
\]

The above lemma, the triangle inequality, Lemma 3.8, and Lemma 3.12 complete the proof of the main result:

**Theorem 3.16.** We have

\[
\| q - q_h \|_{T_h} \lesssim h^k(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}), \\
(3.24a)
\]

\[
\| p - p_h \|_{T_h} \lesssim h^k(|q|_{k+1} + |y|_{k+1} + |p|_{k+1} + |z|_{k+1}). \\
(3.24b)
\]

### 4 Numerical Experiments

Next, we present a numerical example to illustrate our theoretical results. We consider the distributed control problem for the Poisson equation on a square domain \( \Omega = [0, 1] \times [0, 1] \subset \mathbb{R}^2 \) and take \( \gamma = 1 \). We set the exact state and dual state to be \( y(x_1, x_2) = \sin(\pi x_1) \) and \( z(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2) \), and generate the data \( f, g, \) and \( y_d \) from the optimality system (1.3). Numerical results for \( k = 1 \) and \( k = 2 \) for this problem are shown in Table 1–Table 2. The numerical convergence rates match the theory.

### 5 Conclusions

We proposed an EDG method to approximate the solution of an optimal distributed control problems for the Poisson equation. We obtained optimal a priori error estimates for the control, state, and dual state, but suboptimal estimates for their fluxes. As mentioned earlier, EDG has potential
An EDG Method for Distributed Optimal Control of Elliptic PDEs

| $h/\sqrt{2}$ | $1/8$ | $1/16$ | $1/32$ | $1/64$ | $1/128$ |
|------------|------|-------|-------|-------|-------|
| $\|q - q_h\|_{0,\Omega}$ | 2.6598e-02 | 6.7755e-03 | 1.7029e-03 | 4.2631e-04 | 1.0662e-04 |
| order | - | 1.97 | 1.99 | 2.00 | 2.00 |
| $\|p - p_h\|_{0,\Omega}$ | 2.6694e-02 | 7.0861e-03 | 1.7999e-03 | 4.5178e-04 | 1.1306e-04 |
| order | - | 1.91 | 1.98 | 1.99 | 2.00 |
| $\|y - y_h\|_{0,\Omega}$ | 8.3274e-04 | 1.0672e-04 | 1.3592e-05 | 1.7164e-06 | 2.1566e-07 |
| order | - | 2.96 | 2.97 | 2.99 | 3.00 |
| $\|z - z_h\|_{0,\Omega}$ | 1.4515e-03 | 1.9483e-04 | 2.5202e-05 | 3.2009e-06 | 4.0316e-07 |
| order | - | 2.90 | 2.95 | 2.98 | 2.99 |

Table 2: Errors for the state $y$, adjoint state $z$, and the fluxes $q$ and $p$ when $k = 2$.

for optimal control problems involving convection dominated partial differential equations and fluid flows. These problems would be interesting to explore in the future.

Also, we used a different EDG error analysis strategy to prove the error estimates in this work. We are currently investigating another EDG method, and we have used the different analysis approach to prove optimal convergence rates for all variables. The details will be reported in a future paper.

References

[1] M. Braack, Optimal control in fluid mechanics by finite elements with symmetric stabilization, SIAM J. Control Optim., 48 (2009), pp. 672–687.

[2] 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.

[3] Y. Chen, Superconvergence of mixed finite element methods for optimal control problems, Math. Comp., 77 (2008), pp. 1269–1291.

[4] 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.

[5] Y. Chen, L. Dai, and Z. Lu, Superconvergence of rectangular mixed finite element methods for constrained optimal control problem, Adv. Appl. Math. Mech., 2 (2010), pp. 56–75.

[6] Y. Chen, T. Hou, and W. Zheng, Error estimates and superconvergence of mixed finite element methods for optimal control problems with low regularity, Adv. Appl. Math. Mech., 4 (2012), pp. 751–768.

[7] Y. Chen and F. Huang, Galerkin spectral approximation of elliptic optimal control problems with $H^1$-norm state constraint, J. Sci. Comput., 67 (2016), pp. 65–83.

[8] Y. Chen, Y. Huang, W. Liu, and N. Yan, Error estimates and superconvergence of mixed finite element methods for convex optimal control problems, J. Sci. Comput., 42 (2010), pp. 382–403.
[9] Y. Chen, Z. Lu, and Y. Huang, Superconvergence of triangular Raviart-Thomas mixed finite element methods for a bilinear constrained optimal control problem, Comput. Math. Appl., 66 (2013), pp. 1498–1513.

[10] 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.

[11] B. Cockburn, J. Gopalakrishnan, N. C. Nguyen, J. Peraire, and F.-J. Sayas, Analysis of HDG methods for Stokes flow, Math. Comp., 80 (2011), pp. 723–760.

[12] B. Cockburn, J. Guzmán, S.-C. Soon, and H. K. Stolarski, An analysis of the embedded discontinuous Galerkin method for second-order elliptic problems, SIAM J. Numer. Anal., 47 (2009), pp. 2686–2707.

[13] B. Cockburn and K. Mustapha, A hybridizable discontinuous Galerkin method for fractional diffusion problems, Numer. Math., 130 (2015), pp. 293–314.

[14] B. Cockburn and J. Shen, A hybridizable discontinuous Galerkin method for the p-Laplacian, SIAM J. Sci. Comput., 38 (2016), pp. A545–A566.

[15] K. Deckelnick, A. Günther, and M. Hinze, Finite element approximation of elliptic control problems with constraints on the gradient, Numer. Math., 111 (2009), pp. 335–350.

[16] P. Fernandez, N. Nguyen, X. Roca, and J. Peraire, Implicit large-eddy simulation of compressible flows using the interior embedded discontinuous Galerkin method, in 54th AIAA Aerospace Sciences Meeting, AIAA SciTech Forum, AIAA 2016-1332, 2016, https://doi.org/10.2514/6.2016-1332.

[17] 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.

[18] G. Fu and C.-W. Shu, Analysis of an embedded discontinuous galerkin method with implicit-explicit time-marching for convection-diffusion problems, International Journal of Numerical Analysis & Modeling, 14 (2017), pp. 477–499.

[19] D. Givoli, A direct approach to the finite element solution of elliptic optimal control problems, Numer. Methods Partial Differential Equations, 15 (1999), pp. 371–388.

[20] S. Güzey, B. Cockburn, and H. K. Stolarski, The embedded discontinuous Galerkin method: application to linear shell problems, Internat. J. Numer. Methods Engrg., 70 (2007), pp. 757–790.

[21] T. Hou and L. Li, Superconvergence and a posteriori error estimates of splitting positive definite mixed finite element methods for elliptic optimal control problems, Appl. Math. Comput., 273 (2016), pp. 1196–1207.

[22] T. Hou, C. Liu, and Y. Yang, Error estimates and superconvergence of a mixed finite element method for elliptic optimal control problems, Comput. Math. Appl., 74 (2017), pp. 714–726.
An EDG Method for Distributed Optimal Control of Elliptic PDEs

[23] W. Hu, J. Shen, J. R. Singler, Y. Zhang, and X. Zheng, An HDG method for distributed control of convection diffusion PDEs. submitted.

[24] W. Hu, J. Shen, J. R. Singler, Y. Zhang, and X. Zheng, A superconvergent HDG method for distributed control of convection diffusion PDEs. submitted.

[25] 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. submitted.

[26] D. S. Kamenetskiy, On the relation of the embedded discontinuous Galerkin method to the stabilized residual-based finite element methods, Appl. Numer. Math., 108 (2016), pp. 271–285, https://doi.org/10.1016/j.apnum.2016.01.004.

[27] 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.

[28] W. Liu, W. Gong, and N. Yan, A new finite element approximation of a state-constrained optimal control problem, J. Comput. Math., 27 (2009), pp. 97–114.

[29] N. C. Nguyen, J. Peraire, and B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations, J. Comput. Phys., 228 (2009), pp. 3232–3254.

[30] N. C. Nguyen, J. Peraire, and B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for nonlinear convection-diffusion equations, J. Comput. Phys., 228 (2009), pp. 8841–8855.

[31] N. C. Nguyen, J. Peraire, and B. Cockburn, A hybridizable discontinuous Galerkin method for Stokes flow, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 582–597.

[32] N. C. Nguyen, J. Peraire, and B. Cockburn, A class of embedded discontinuous Galerkin methods for computational fluid dynamics, J. Comput. Phys., 302 (2015), pp. 674–692.

[33] J. Peraire, N. Nguyen, and B. Cockburn, An embedded discontinuous Galerkin method for the compressible Euler and Navier-Stokes equations, in Proceedings of the 20th AIAA Computational Fluid Dynamics Conference, AIAA 2011-3228, 2011, https://doi.org/10.2514/6.2011-3228.

[34] M. Stanglmeier, N. C. Nguyen, J. Peraire, and B. Cockburn, An explicit hybridizable discontinuous Galerkin method for the acoustic wave equation, Comput. Methods Appl. Mech. Engrg., 300 (2016), pp. 748–769.

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

[36] H. Zhu and F. Celiker, Error analysis of an HDG method for a distributed optimal control problem, J. Comput. Appl. Math., 307 (2016), pp. 2–12.