Numerische Mathematik (2020) 144:375–411
https://doi.org/10.1007/s00211-019-01090-2

Numerische Mathematik

A superconvergent hybridizable discontinuous Galerkin method for Dirichlet boundary control of elliptic PDEs

Weiwei Hu[1] · Jiguang Shen[2] · John R. Singler[3] · Yangwen Zhang[3] · Xiaobo Zheng[4]

Received: 7 December 2017 / Revised: 12 July 2019 / Published online: 11 December 2019

Abstract
We begin an investigation of hybridizable discontinuous Galerkin (HDG) methods for approximating the solution of Dirichlet boundary control problems governed by elliptic PDEs. These problems can involve atypical variational formulations, and often have solutions with low regularity on polyhedral domains. These issues can provide challenges for numerical methods and the associated numerical analysis. We propose an HDG method for a Dirichlet boundary control problem for the Poisson equation, and obtain optimal a priori error estimates for the control. Specifically, under certain assumptions, for a 2D convex polygonal domain we show the control converges at a superlinear rate. We present 2D and 3D numerical experiments to demonstrate our theoretical results.

Mathematics Subject Classification 65N30 · 49M25

✉ John R. Singler
singlerj@mst.edu

Weiwei Hu
weiwei.hu@okstate.edu

Jiguang Shen
shenx179@umn.edu

Yangwen Zhang
ywzfg4@mst.edu

Xiaobo Zheng
zhengxiaobosc@yahoo.com

1 Department of Mathematics, Oklahoma State University, Stillwater, OK, USA

2 School of Mathematics, University of Minnesota, Minneapolis, MN, USA

3 Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO, USA

4 College of Mathematics, Sichuan University, Chengdu, China

1 Introduction

We consider the following elliptic Dirichlet boundary control problem on a Lipschitz polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with boundary $\Gamma = \partial\Omega$:

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

where $\gamma > 0$ and y is the solution of the Poisson equation with nonhomogeneous Dirichlet boundary conditions

$$-\Delta y = f \quad \text{in } \Omega, \tag{2}$$
$$y = u \quad \text{on } \Gamma. \tag{3}$$

It is well known that the Dirichlet boundary control problem (1)–(3) is equivalent to the optimality system

$$-\Delta y = f \qquad \text{in } \Omega, \tag{4a}$$
$$y = u \qquad \text{on } \Gamma, \tag{4b}$$
$$-\Delta z = y - y_d \quad \text{in } \Omega, \tag{4c}$$
$$z = 0 \qquad \text{on } \Gamma, \tag{4d}$$
$$u = \gamma^{-1}\frac{\partial z}{\partial \boldsymbol{n}} \quad \text{on } \Gamma, \tag{4e}$$

where $\boldsymbol{n}$ is the unit outer normal to Γ.

Dirichlet boundary control has many applications in fluid flow problems and other fields, and therefore the mathematical study of these control problems has become an important area of research. Major theoretical and computational developments have been made in the recent past; see, e.g., [8,19,20,24–26,31–34,53,56,58]. However, only in the last 10 years have researchers developed thorough well-posedness, regularity, and finite element error analysis results for elliptic PDEs; see [1,6,21,42,59] and the references therein. One difficulty of Dirichlet boundary control problems is that the Dirichlet boundary data does not directly enter a standard variational setting for the PDE; instead, the state equation is understood in a very weak sense. Also, solutions of the optimality system typically do not have high regularity on polyhedral domains; corners cause the normal derivative of the adjoint state $\partial z/\partial \boldsymbol{n}$ in the optimality condition (4) to have limited smoothness. Solutions with limited regularity can lead to complications for numerical methods and numerical analysis.

To avoid the difficulties described above, researchers have considered other approaches including modified cost functionals [13,30,32,48], approximating the Dirichlet boundary condition with a Robin boundary condition [2,3,5,35,53], and weak boundary penalization [9].

One way to approximate the solution of the original problem without penalization and also avoid the variational difficulty is to use a mixed finite element method. Recently, Gong and Yan [29] considered this approach and obtained

$$\|u - u_h\|_{0,\Gamma} = O(h^{1-1/s})$$

when the control belongs to $H^{1-1/s}(\Gamma)$ and the lowest order Raviart–Thomas elements are used for the computation.

As researchers continue to investigate Dirichlet boundary control problems of increasingly complexity, it may become natural to utilize discontinuous Galerkin methods for the spatial discretization of problems involving strong convection and discontinuities. We have performed preliminary computations using an hybridizable discontinuous Galerkin (HDG) method for a similar Dirichlet boundary control problem for the Stokes equations. Our preliminary results for this problem indicate that the optimal control can indeed be discontinuous at the corners of the domain. Before we continue to investigate problems of such complexity, we begin this line of research by considering an HDG method to approximate the solution of the above Dirichlet boundary control problem.

HDG methods also utilize a mixed formulation and therefore avoid the variational difficulty of the Dirichlet control problem. Furthermore, the number of degrees of freedom for HDG methods are much less than standard mixed methods or other DG approaches. Moreover, the RT element is a special case of the HDG method. We provide more background about HDG methods in Sect. 3.

We propose an HDG method to approximate the control in Sect. 3, and in Sect. 4 we prove an optimal superlinear rate of convergence for the control in 2D using discontinuous linear elements for the control on a quasi-uniform mesh under certain assumptions on the domain and y_d. To give a specific example, for a rectangular 2D domain and $y_d \in H^1(\Omega) \cap L^\infty(\Omega)$, we obtain the following a priori error bounds for the state y, adjoint state z, their fluxes $\boldsymbol{q} = -\nabla y$ and $\boldsymbol{p} = -\nabla z$, and the optimal control u:

$$\begin{aligned} \|y - y_h\|_{0,\Omega} &= O(h^{3/2-\varepsilon}), & \|z - z_h\|_{0,\Omega} &= O(h^{3/2-\varepsilon}), \\ \|\boldsymbol{q} - \boldsymbol{q}_h\|_{0,\Omega} &= O(h^{1-\varepsilon}), & \|\boldsymbol{p} - \boldsymbol{p}_h\|_{0,\Omega} &= O(h^{3/2-\varepsilon}), \end{aligned}$$

and

$$\|u - u_h\|_{0,\Gamma} = O(h^{3/2-\varepsilon}),$$

for any $\varepsilon > 0$. We demonstrate the performance of the HDG method with numerical experiments in 2D and 3D in Sect. 5.

We note that an optimal superlinear convergence rate for the control on polygonal domains has also been recently obtained in [1, Theorem 4.1 and Remark 4.8] for standard continuous finite elements using linear elements on a superconvergence mesh or quadratic elements on a quasi-uniform mesh. (Also see [21] for similar results on curved domains.) The number of degrees of freedom for these standard finite element methods is lower than the HDG method considered here with discontinuous linear elements for the control. However, we do not give a thorough comparison of the methods here since, as mentioned above, the primary goal of this work is to begin an investigation of HDG methods for Dirichlet boundary control problems before moving to more complex problems, for which DG methods may be more appropriate.

2 Background: the optimality system and regularity

To begin, we review some fundamental results concerning the optimality system for the control problem in polyhedral domains and the regularity of the solution in 2D polygonal domains. We also provide a mixed formulation of the optimality system and prove it is well-posed.

Throughout the paper we adopt the standard notation $W^{m,p}(\Omega)$ for Sobolev spaces on Ω 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 $|\cdot|_{m,\Omega}$. Also, $H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$. We denote the L^2-inner products on $L^2(\Omega)$ and $L^2(\Gamma)$ by

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

Define the space $H(\text{div}; \Omega)$ as

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

The inner product $(\cdot, \cdot)$ is defined above only for scalar variables, but we use the same inner product notation for vector functions: $(\boldsymbol{q}, \boldsymbol{p}) = \int_\Omega \boldsymbol{q}\cdot\boldsymbol{p}$ for all $\boldsymbol{q}, \boldsymbol{p} \in [L^2(\Omega)]^d$. Throughout this paper, the norm corresponding to a given inner product is defined in the standard way.

We use the bracket $[\cdot, \cdot]_\Gamma$ to denote the $H^{-1/2}(\Gamma)\times H^{1/2}(\Gamma)$ duality pairing. Recall for $\boldsymbol{r} \in H(\text{div}, \Omega)$, it is known that $\boldsymbol{r}\cdot\boldsymbol{n} \in H^{-1/2}(\Gamma)$. Also, for any $v \in H^1(\Omega)$ with boundary trace $v|_\Gamma \in H^{1/2}(\Gamma)$, we have the integration by parts formula (see, e.g., [55, Theorem 6.1])

$$(v, \nabla\cdot\boldsymbol{r}) = [v|_\Gamma, \boldsymbol{r}\cdot\boldsymbol{n}]_\Gamma - (\nabla v, \boldsymbol{r}) \quad \text{for all } \boldsymbol{r} \in H(\text{div}, \Omega). \tag{5}$$

Furthermore, for any $v_\Gamma \in H^{1/2}(\Gamma)$ and $\boldsymbol{r} \in H(\text{div}, \Omega)$ satisfying $\boldsymbol{r}\cdot\boldsymbol{n} \in L^2(\Gamma)$, we have that the duality pairing reduces to the $L^2(\Gamma)$ inner product (see, e.g., [55, Section 6.1]), i.e.,

$$[v|_\Gamma, \boldsymbol{r}\cdot\boldsymbol{n}]_\Gamma = \langle v|_\Gamma, \boldsymbol{r}\cdot\boldsymbol{n}\rangle.$$

2.1 The optimality system

For the remainder of this section, we assume $f, y_d \in L^2(\Omega)$ and $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded convex polyhedral domain with boundary Γ. Below, we recall a precise well-posedness result from [42] for the optimal control problem.

Lemma 2.1 [42, Lemma 2.6] *The pair* $(u, y) \in L^2(\partial\Omega) \times L^2(\Omega)$ *is a solution of the optimal control problem if and only if there exists an adjoint state* $z \in H^2(\Omega) \cap H_0^1(\Omega)$ *such that* (u, y, z) *solves the optimality system*

$$-(y, \Delta\varphi) + \langle u, \partial\varphi/\partial\boldsymbol{n}\rangle = (f, \varphi) \qquad \forall\varphi \in H^2 \cap H_0^1(\Omega), \tag{6a}$$

$$-(\psi, \Delta z) - (y, \psi) = -(y_d, \psi) \qquad \forall\psi \in L^2(\Omega), \tag{6b}$$

$$\langle \gamma u - \partial z/\partial\boldsymbol{n}, \chi\rangle = 0 \qquad \forall\chi \in L^2(\partial\Omega). \tag{6c}$$

Remark 2.2 In [42], the authors only consider the case of a convex polygonal domain, but their proof also works for a convex polyhedral domain in $\mathbb{R}^3$.

As mentioned earlier, the state Eqs. (2), (3) in the optimal control problem is written formally and should be understood to hold in the very weak sense (6a). Similarly, the formal optimality system (4) should be understood in the weak sense (6).

The convexity of the domain gives additional regularity for u and y.

Lemma 2.3 *The solution of the optimality system satisfies*

$$y \in H^1(\Omega), \quad \partial z/\partial\boldsymbol{n} \in H^{1/2}(\Gamma), \quad u \in H^{1/2}(\Gamma).$$

Proof Since $y - y_d \in L^2(\Omega)$ and Ω is a convex polyhedral domain, Eq. (6b) along with [5, Lemma A.2] for the 2D case and [28, Equation (3.5) with $q = 2$] for the 3D case give $\partial z/\partial\boldsymbol{n} \in H^{1/2}(\Gamma)$. The desired regularity for u follows from (6c). Lastly, Eq. (6a), $f \in L^2(\Omega)$, and $u \in H^{1/2}(\Gamma)$ imply $y \in H^1(\Omega)$. □

To avoid the variational difficulty we follow the strategy introduced by Wei Gong and Ningning Yan [29] and consider a mixed formulation of the optimality system. Introduce two flux variables $\boldsymbol{q} = -\nabla y$ and $\boldsymbol{p} = -\nabla z$. We consider the following mixed weak form of (4a)-(4e): Find $(\boldsymbol{q}, y, \boldsymbol{p}, z, u) \in H(\text{div}, \Omega) \times L^2(\Omega) \times H(\text{div}, \Omega) \times L^2(\Omega) \times H^{1/2}(\Gamma)$ such that

$$(\boldsymbol{q}, \boldsymbol{r}_1) - (y, \nabla\cdot\boldsymbol{r}_1) + [u, \boldsymbol{r}_1\cdot\boldsymbol{n}]_\Gamma = 0, \tag{7a}$$

$$(\nabla\cdot\boldsymbol{q}, w_1) = (f, w_1), \tag{7b}$$

$$(\boldsymbol{p}, \boldsymbol{r}_2) - (z, \nabla\cdot\boldsymbol{r}_2) = 0, \tag{7c}$$

$$(\nabla\cdot\boldsymbol{p}, w_2) - (y, w_2) = -(y_d, w_2), \tag{7d}$$

$$[\gamma u + \boldsymbol{p}\cdot\boldsymbol{n}, \xi]_\Gamma = 0 \tag{7e}$$

for all $(\boldsymbol{r}_1, w_1, \boldsymbol{r}_2, w_2, \xi) \in H(\text{div}, \Omega) \times L^2(\Omega) \times H(\text{div}, \Omega) \times L^2(\Omega) \times H^{1/2}(\Gamma)$. This mixed formulation can be derived directly from the formally stated optimality system (4). Below, we work with the precise statement of the optimality system (6) and prove the well-posedness of this mixed problem.

Since we require $u \in H^{1/2}(\Gamma)$, all of the duality pairings in the above mixed formulation are well defined. Since the domain is convex, Lemma 2.3 implies this requirement for the optimal control u is not a restriction.

Next, we show the well-posedness of the mixed weak form of the optimality system.

Theorem 2.4 *Let* (u, y, z) *be the solution of the optimality system* (6)*, and define* $\boldsymbol{q} = -\nabla y$ *and* $\boldsymbol{p} = -\nabla z$*. Then* $(\boldsymbol{q}, y, \boldsymbol{p}, z, u) \in H(\mathrm{div}, \Omega) \times L^2(\Omega) \times H(\mathrm{div}, \Omega) \times L^2(\Omega) \times H^{1/2}(\Gamma)$ *is the unique solution of the mixed optimality system* (7)*.*

Proof Let $(u, y, z) \in L^2(\Gamma) \times L^2(\Omega) \times [H^2(\Omega) \cap H_0^1(\Omega)]$ be the unique solution of the optimality system (6). Lemma 2.3 gives the additional regularity $u \in H^{1/2}(\Gamma)$, $y \in H^1(\Omega)$, and $\partial z/\partial \boldsymbol{n} \in H^{1/2}(\Gamma)$. Furthermore, this regularity along with (6a) gives that u is the boundary trace of y. Then $(\boldsymbol{q}, y, \boldsymbol{p}, z, u)$ give a solution to the mixed formulation. To be complete, we give the details.

For $\boldsymbol{r}_1, \boldsymbol{r}_2 \in H(\mathrm{div}, \Omega)$, Eqs. (7a) and (7c) follow directly from the integration by parts formula (5) applied to $y \in H^1(\Omega)$ and $z \in H_0^1(\Omega)$, respectively. Also, since $\partial z/\partial \boldsymbol{n} \in H^{1/2}(\Gamma)$, $\boldsymbol{p} = -\nabla z$, and $u = -\gamma^{-1}\partial z/\partial \boldsymbol{n}$ holds in $L^2(\Gamma)$, it is clear that (7e) holds. Furthermore, since $z \in H^2 \cap H_0^1(\Omega)$ satisfies the dual problem (6b) and $\boldsymbol{p} = -\nabla z \in H(\mathrm{div}, \Omega)$, it is clear that (7d) holds.

It remains to show $\boldsymbol{q} \in H(\mathrm{div}, \Omega)$ and that (7b) holds. Since $u \in H^{1/2}(\Gamma)$, $y \in H^1(\Omega)$, u is the boundary trace of y, and $\boldsymbol{q} = -\nabla y$, the very weak form of the state Eq. (6a) gives that for any $\varphi \in H^2 \cap H_0^1(\Omega)$ we have

$$(f, \varphi) = -(y, \Delta\varphi) + \langle u, \partial\varphi/\partial \boldsymbol{n}\rangle = -(\boldsymbol{q}, \nabla\varphi).$$

Restricting to $\varphi \in C_0^\infty(\Omega)$ gives, by definition, that $\nabla \cdot \boldsymbol{q} = f \in L^2(\Omega)$ in the distributional sense, i.e.,

$$(f, \varphi) = (\nabla \cdot \boldsymbol{q}, \varphi) \quad \text{for all } \varphi \in C_0^\infty(\Omega).$$

This implies $\boldsymbol{q} \in H(\mathrm{div}, \Omega)$. Also, by density, the above equation holds for all $\varphi \in L^2(\Omega)$, and therefore (7b) holds.

Next, to prove the uniqueness we set $f = y_d = 0$ and show zero is the only solution of the mixed formulation (7). Take $(\boldsymbol{r}_1, w_1) = (\boldsymbol{p}, -z)$, $(\boldsymbol{r}_2, w_2) = (-\boldsymbol{q}, y)$, and $\xi = -u$, and then sum the equations to give

$$\gamma \|u\|^2_{L^2(\Gamma)} + \|y\|^2_{L^2(\Omega)} = 0.$$

Since $\gamma > 0$, this implies $u = 0$ and $y = 0$. Next, since $y = 0$ and $u = 0$, taking $\boldsymbol{r}_1 = \boldsymbol{q}$ implies $\boldsymbol{q} = \boldsymbol{0}$. Also, taking $\boldsymbol{r}_2 = \boldsymbol{p}$ and $w_2 = z$ and summing these two equations implies $\boldsymbol{p} = \boldsymbol{0}$. This leaves

$$0 = -(z, \nabla \cdot \boldsymbol{r}_2) \quad \text{for all } \boldsymbol{r}_2 \in H(\mathrm{div}, \Omega).$$

Let z_2 be the solution of $-\Delta z_2 = z$ with zero Dirichlet boundary conditions. Since the domain is convex, we have $z_2 \in H^2(\Omega) \cap H_0^1(\Omega)$. Set $\boldsymbol{r}_2 = \nabla z_2 \in H(\mathrm{div}, \Omega)$. This gives $0 = \|z\|^2_{L^2(\Omega)}$, which implies $z = 0$. This proves the uniqueness. □

2.2 Regularity

One of the main reasons that the Dirichlet boundary control problem can be challenging numerically is that the solution can have very low regularity, and this restricts the convergence rates of finite element and DG methods. In order to prove a superlinear convergence rate for the optimal control for the HDG method in Sect. 4, we assume the solution has the following fractional Sobolev regularity:

$$u \in H^{r_u}(\Gamma), \quad y \in H^{r_y}(\Omega), \quad z \in H^{r_z}(\Omega), \quad \boldsymbol{q} \in H^{r_{\boldsymbol{q}}}(\Omega), \quad \boldsymbol{p} \in H^{r_{\boldsymbol{p}}}(\Omega), \tag{8}$$

with

$$r_u > 1, \quad r_y > 1, \quad r_z > 2, \quad r_{\boldsymbol{q}} > 1/2, \quad r_{\boldsymbol{p}} > 1. \tag{9}$$

We require $r_{\boldsymbol{q}} > 1/2$ in order to guarantee q has a well-defined trace on the boundary Γ. We note that it may be possible to use the techniques in [38] to lower the regularity requirement on $\boldsymbol{q}$. We leave this to be considered elsewhere.

For a 2D convex polygonal domain and $f = 0$, we use a recent regularity result of Mateos and Neitzel [41] below to give conditions on the domain and y_d to guarantee the solution has the above regularity. For a higher dimensional convex polyhedral domain, the regularity theory is much more complicated, and we do not attempt to provide conditions to guarantee the above regularity in this work.

Theorem 2.5 [41, Lemma 3 and Corollary 1] *Assume Ω is a convex polygonal domain and $f = 0$. Let $\omega \in [\pi/3, \pi)$ be the largest interior angle of Γ, and define p_Ω, r_Ω by*

$$p_\Omega = \frac{2}{2 - \pi/\max\{\omega, \pi/2\}} \in (2, \infty],$$

and

$$r_\Omega = 1 + \frac{\pi}{\omega} \in (2, 4].$$

If $y_d \in L^p(\Omega) \cap H^{r-2}(\Omega)$ for all $p < p_\Omega$ and $r < r_\Omega$, then the solution (u, y, z) satisfies

$$\begin{aligned} u &\in H^{r-3/2}(\Gamma) \cap W^{1-1/p,p}(\Gamma), \\ y &\in H^{r-1}(\Omega) \cap W^{1,p}(\Omega), \\ z &\in H_0^1(\Omega) \cap H^r(\Omega) \cap W^{2,p}(\Omega) \end{aligned}$$

for all

$$p < p_\Omega, \quad r < \min\{3, r_\Omega\}.$$

Corollary 2.6 *Under the assumptions of Theorem 2.5, the flux variables $\boldsymbol{q} = -\nabla y$ and $\boldsymbol{p} = -\nabla z$ satisfy*

$$\boldsymbol{q} \in H^{r-2}(\Omega) \cap H(\mathrm{div}, \Omega), \quad \boldsymbol{p} \in H^{r-1}(\Omega) \cap H(\mathrm{div}, \Omega)$$

for all $r < \min\{3, r_\Omega\}$.

The regularity for the flux variable $\boldsymbol{q} = -\nabla y$ is low; Corollary 2.6 only guarantees $\boldsymbol{q} \in H^{r_q}$ for some $0 < r_{\boldsymbol{q}} < 1$. For the HDG approximation theory, we need the regularity condition $r_{\boldsymbol{q}} > 1/2$. We can guarantee this condition by restricting the maximum interior angle ω. Specifically, if y_d has the required smoothness and ω satisfies

$$\omega \in [\pi/3, 2\pi/3),$$

then $r_\Omega \in (5/2, 4]$ and we are guaranteed $\boldsymbol{q} \in H^{r_q}$ for some $r_{\boldsymbol{q}} > 1/2$.

Also, when we restrict $\omega \in [\pi/3, 2\pi/3)$ as above, this guarantees $u \in H^{r_u}$ for some $1 < r_u < 3/2$ and furthermore the regularity assumption (8), (9) is satisfied. For a rectangular domain, we have $p_\Omega = \infty$ and $r_\Omega = 3$. Therefore if $y_d \in H^1(\Omega) \cap L^\infty(\Omega)$ we are guaranteed the fractional Sobolev regularity

$$r_u = \frac{3}{2} - \varepsilon, \quad r_y = 2 - \varepsilon, \quad r_z = 3 - \varepsilon, \quad r_{\boldsymbol{q}} = 1 - \varepsilon, \quad r_{\boldsymbol{p}} = 2 - \varepsilon$$

for any $\varepsilon > 0$.

3 HDG formulation and implementation

A mixed method can avoid the variational difficulty by introducing flux variables $\boldsymbol{q}$ and $\boldsymbol{p}$ and the equation for the optimal control (7e). However, these two additional vector variables will increase the computational cost, even if the lowest order RT method is used.

We introduce an HDG method for the optimality system (4) to take advantage of the mixed formulation and also reduce the computational cost compared to standard mixed methods. Specifically, we introduce the flux variables but eliminate them before we solve the global equation; this significantly reduces the number of degrees of freedom.

HDG methods were proposed by Cockburn et al. in [15] as an improvement of tradition discontinuous Galerkin methods and have many applications; see, e.g., [7,11,16–18,44–47,57]. The approximate scalar variable and flux are expressed in an element-by-element fashion in terms of an approximate trace of the scalar variable along the element boundary. Then, a unique value for the trace at inter-element boundaries is obtained by enforcing flux continuity. This leads to a global equation system in terms of the approximate boundary traces only. The high number of globally coupled degrees of freedom is significantly reduced compared to other DG methods and standard mixed methods.

Before we introduce the HDG method, we first set some notation. Let $\{\mathcal{T}_h\}$ be a conforming quasi-uniform polyhedral mesh of Ω. We denote by $\partial\mathcal{T}_h$ the set $\{\partial K : K \in \mathcal{T}_h\}$. For an element K of the collection $\mathcal{T}_h$, 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^+ and

K^- of the collection $\mathcal{T}_h$, 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 ε_h^o and ε_h^∂ denote the sets of interior and boundary faces, respectively. We denote by ε_h the union of ε_h^o and ε_h^∂. We finally introduce

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

Similar to the notation defined in Sect. 2, we also use the inner product notation $(\cdot,\cdot)_{\mathcal{T}_h}$ for vector functions.

Let $\mathcal{P}^k(D)$ denote the set of polynomials of degree at most k on a domain D. We introduce the discontinuous finite element spaces

$$\boldsymbol{V}_h := \{\boldsymbol{v} \in [L^2(\Omega)]^d : \boldsymbol{v}|_K \in [\mathcal{P}^k(K)]^d, \forall K \in \mathcal{T}_h\}, \tag{10}$$

$$W_h := \{w \in L^2(\Omega) : w|_K \in \mathcal{P}^{k+1}(K), \forall K \in \mathcal{T}_h\}, \tag{11}$$

$$M_h := \{\mu \in L^2(\varepsilon_h) : \mu|_e \in \mathcal{P}^k(e), \forall e \in \varepsilon_h\}. \tag{12}$$

The space W_h is for scalar variables, while $\boldsymbol{V}_h$ is for flux variables and M_h is for boundary trace variables. Note that the polynomial degree for the scalar variables is one order higher than the polynomial degree for the other variables. Also, the boundary trace variables will be used to eliminate the state and flux variables from the coupled global equations, thus substantially reducing the number of degrees of freedom.

Let $M_h(o)$ and $M_h(\partial)$ denote the spaces defined in the same way as M_h, but with ε_h replaced by ε_h^o and ε_h^∂, respectively. Note that M_h consists of functions which are continuous inside the faces (or edges) $e \in \varepsilon_h$ and discontinuous at their borders. In addition, for any function $w \in W_h$ we use ∇w to denote the piecewise gradient on each element $K \in \mathcal{T}_h$. A similar convention applies to the divergence operator $\nabla\cdot\boldsymbol{r}$ for all $\boldsymbol{r} \in \boldsymbol{V}_h$.

3.1 The HDG formulation

To approximate the solution of the mixed weak form (4a)–(4e) of the optimality system, the HDG method seeks approximate fluxes $\boldsymbol{q}_h, \boldsymbol{p}_h \in \boldsymbol{V}_h$, states $y_h, z_h \in W_h$, interior element boundary traces $\widehat{y}_h^o, \widehat{z}_h^o \in M_h(o)$, and boundary control $u_h \in M_h(\partial)$ satisfying

$$(\boldsymbol{q}_h, \boldsymbol{r_1})_{\mathcal{T}_h} - (y_h, \nabla\cdot\boldsymbol{r_1})_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{r_1}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle u_h, \boldsymbol{r_1}\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} = 0, \tag{13a}$$

$$-(\boldsymbol{q}_h, \nabla w_1)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}, w_1\rangle_{\partial\mathcal{T}_h} = (f, w_1)_{\mathcal{T}_h} \tag{13b}$$

for all $(\boldsymbol{r_1}, w_1) \in \boldsymbol{V}_h \times W_h$,

$$(\boldsymbol{p}_h, \boldsymbol{r_2})_{\mathcal{T}_h} - (z_h, \nabla\cdot\boldsymbol{r_2})_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{r_2}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0, \tag{13c}$$

$$-(\boldsymbol{p}_h, \nabla w_2)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}, w_2 \rangle_{\partial \mathcal{T}_h} - (y_h, w_2)_{\mathcal{T}_h} = -(y_d, w_2)_{\mathcal{T}_h} \tag{13d}$$

for all $(\boldsymbol{r_2}, w_2) \in \boldsymbol{V}_h \times W_h$,

$$\langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} = 0 \tag{13e}$$

for all $\mu_1 \in M_h(o)$,

$$\langle \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} = 0 \tag{13f}$$

for all $\mu_2 \in M_h(o)$, and

$$\langle u_h, \mu_3 \rangle_{\varepsilon_h^\partial} + \langle \gamma^{-1} \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}, \mu_3 \rangle_{\varepsilon_h^\partial} = 0 \tag{13g}$$

for all $\mu_3 \in M_h(\partial)$.

The numerical traces on $\partial \mathcal{T}_h$ are defined as

$$\widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n} = \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1}(P_M y_h - \widehat{y}_h^o) \quad \text{on } \partial \mathcal{T}_h \backslash \varepsilon_h^\partial, \tag{13h}$$

$$\widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n} = \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1}(P_M y_h - u_h) \quad \text{on } \varepsilon_h^\partial, \tag{13i}$$

$$\widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n} = \boldsymbol{p}_h \cdot \boldsymbol{n} + h^{-1}(P_M z_h - \widehat{z}_h^o) \quad \text{on } \partial \mathcal{T}_h \backslash \varepsilon_h^\partial, \tag{13j}$$

$$\widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n} = \boldsymbol{p}_h \cdot \boldsymbol{n} + h^{-1} P_M z_h \qquad \text{on } \varepsilon_h^\partial, \tag{13k}$$

where P_M denotes the standard L^2-orthogonal projection from $L^2(\varepsilon_h)$ onto M_h. This completes the formulation of the HDG method.

The HDG formulation with h^{-1} stabilization, polynomial degree $k+1$ for the scalar unknown, and polynomial degree k for the other unknowns was originally introduced by Lehrenfeld in [36] and first analyzed by Oikawa in [49]. See [14, Sections 6.5-6.6] for more about the history of this method, and its relationship to other HDG and DG methods. This HDG method is widely considered to be a superconvergent method. Specifically, if the solution of the PDE is smooth enough, then in many cases $O(h^{k+2})$ error estimates can be obtained for the state variable for all $k \geq 0$; see [38,49] for the Poisson equation, [50] for linear elasticity, [51] for linear convection diffusion, and [52] the Navier–Stokes equations. Hence, from the viewpoint of globally coupled degrees of freedom, this method achieves superconvergence for the scalar variable. However, this method loses the superconvergence when $k = 0$ for convection diffusion [51] and the Navier–Stokes equations [52]. A fix for linear convection diffusion is found in [10], but a fix for the Navier–Stokes equations is not known.

3.2 Implementation

To arrive at the HDG formulation we implement numerically, we insert (13h)–(13k) into (13a)–(13g), and find after some simple manipulations that

$$(\boldsymbol{q}_h, \boldsymbol{p}_h, y_h, z_h, \widehat{y}_h^o, \widehat{z}_h^o, u_h) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o) \times M_h(\partial)$$

is the solution of the following weak formulation:

$$(\boldsymbol{q}_h, \boldsymbol{r_1})_{\mathcal{T}_h} - (y_h, \nabla \cdot \boldsymbol{r_1})_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{r_1} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle u_h, \boldsymbol{r_1} \cdot \boldsymbol{n} \rangle_{\varepsilon_h^\partial} = 0, \tag{14a}$$

$$(\boldsymbol{p}_h, \boldsymbol{r_2})_{\mathcal{T}_h} - (z_h, \nabla \cdot \boldsymbol{r_2})_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{r_2} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} = 0, \tag{14b}$$

$$(\nabla \cdot \boldsymbol{q}_h, w_1)_{\mathcal{T}_h} + \langle h^{-1} P_M y_h, w_1 \rangle_{\partial \mathcal{T}_h} - \langle h^{-1} \widehat{y}_h^o, w_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} \tag{14c}$$

$$-\langle h^{-1} u_h, w_1 \rangle_{\varepsilon_h^\partial} = (f, w_1)_{\mathcal{T}_h}, \tag{14d}$$

$$(\nabla \cdot \boldsymbol{p}_h, w_2)_{\mathcal{T}_h} + \langle h^{-1} P_M z_h, w_2 \rangle_{\partial \mathcal{T}_h} - \langle h^{-1} \widehat{z}_h^o, w_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} \tag{14e}$$

$$-(y_h, w_2)_{\mathcal{T}_h} = -(y_d, w_2)_{\mathcal{T}_h}, \tag{14f}$$

$$\langle \boldsymbol{q}_h \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle h^{-1} y_h, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} - \langle h^{-1} \widehat{y}_h^o, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} = 0, \tag{14g}$$

$$\langle \boldsymbol{p}_h \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle h^{-1} z_h, \mu_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} - \langle h^{-1} \widehat{z}_h^o, \mu_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} = 0, \tag{14h}$$

$$\langle u_h, \mu_3 \rangle_{\varepsilon_h^\partial} + \langle \gamma^{-1} \boldsymbol{p}_h \cdot \boldsymbol{n}, \mu_3 \rangle_{\varepsilon_h^\partial} + \langle \gamma^{-1} h^{-1} z_h, \mu_3 \rangle_{\varepsilon_h^\partial} = 0 \tag{14i}$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2, w_1, w_2, \mu_1, \mu_2, \mu_3) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o) \times M_h(\partial)$.

3.2.1 Matrix equations

Assume $\boldsymbol{V}_h = \mathrm{span}\{\boldsymbol{\varphi}_i\}_{i=1}^{N_1}$, $W_h = \mathrm{span}\{\phi_i\}_{i=1}^{N_2}$, $M_h(o) = \mathrm{span}\{\psi_i\}_{i=1}^{N_3}$, and $M_h(\partial) = \mathrm{span}\{\psi_i\}_{i=1+N_3}^{N_4}$. Then

$$\begin{aligned} \boldsymbol{q}_h &= \sum_{j=1}^{N_1} q_j \boldsymbol{\varphi}_j, \quad \boldsymbol{p}_h = \sum_{j=1}^{N_1} p_j \boldsymbol{\varphi}_j, \quad y_h = \sum_{j=1}^{N_2} y_j \phi_j, \quad z_h = \sum_{j=1}^{N_2} z_j \phi_j, \\ \widehat{y}_h^o &= \sum_{j=1}^{N_3} \alpha_j \psi_j, \quad \widehat{z}_h^o = \sum_{j=1}^{N_3} \gamma_j \psi_j, \quad u_h = \sum_{j=1+N_3}^{N_4} \beta_j \psi_j. \end{aligned} \tag{15}$$

Substitute (15) into (14a)–(14i) and use the corresponding test functions to test (14a)–(14i), respectively, to obtain the matrix equation

$$\begin{bmatrix} A_1 & 0 & -A_2 & 0 & A_8 & 0 & A_9 \\ 0 & A_1 & 0 & -A_2 & 0 & A_8 & 0 \\ A_2^T & 0 & A_5 & 0 & -A_{10} & 0 & -A_{11} \\ 0 & A_2^T & -A_4 & A_5 & 0 & -A_{10} & 0 \\ A_8^T & 0 & A_{10}^T & 0 & A_{11} & 0 & 0 \\ 0 & A_8^T & 0 & A_{10}^T & 0 & A_{11} & 0 \\ 0 & \gamma^{-1} A_{12} & 0 & \gamma^{-1} A_{13} & 0 & 0 & A_{14} \end{bmatrix} \begin{bmatrix} \mathfrak{q} \\ \mathfrak{p} \\ \mathfrak{y} \\ \mathfrak{z} \\ \widehat{\mathfrak{y}} \\ \widehat{\mathfrak{z}} \\ \mathfrak{u} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ b_1 \\ -b_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{16}$$

Here, $\mathfrak{q}$, $\mathfrak{p}$, $\mathfrak{y}$, $\mathfrak{z}$, $\widehat{\mathfrak{y}}$, $\widehat{\mathfrak{z}}$, $\mathfrak{u}$ are the coefficient vectors for $\boldsymbol{q}_h$, $\boldsymbol{p}_h$, y_h, z_h, $\widehat{y}_h^o$, $\widehat{z}_h^o$, u_h, respectively, and

$$
\begin{aligned}
&A_1 = [(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i)_{\mathcal{T}_h}], \quad A_2 = [(\phi_j, \nabla \cdot \boldsymbol{\varphi_i})_{\mathcal{T}_h}], \quad A_3 = [(\psi_j, \boldsymbol{\varphi}_i \cdot \boldsymbol{n})_{\mathcal{T}_h}],\\
&A_4 = [(\phi_j, \phi_i)_{\mathcal{T}_h}], \quad A_5 = [\langle h^{-1} P_M \phi_j, \phi_i\rangle_{\partial\mathcal{T}_h}], \quad A_6 = [\langle h^{-1}\psi_j, \psi_i\rangle_{\partial\mathcal{T}_h}],\\
&A_7 = [\langle h^{-1}\psi_j, \varphi_i\rangle_{\partial\mathcal{T}_h}], \quad b_1 = [(f, \phi_i)_{\mathcal{T}_h}], \quad b_2 = [(y_d, \phi_i)_{\mathcal{T}_h}].
\end{aligned}
$$

The remaining matrices $A_8 - A_{14}$ are constructed by extracting the corresponding rows and columns from A_3, A_6, and A_7. In the actual computation, to save memory we do not assemble the large matrix in Eq. (16).

Equation (16) can be rewritten as

$$
\begin{bmatrix} B_1 & B_2 & B_3 \\ -B_2^T & B_4 & B_5 \\ B_6 & B_7 & B_8 \end{bmatrix}
\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{bmatrix}
=
\begin{bmatrix} 0 \\ \boldsymbol{b} \\ 0 \end{bmatrix}, \tag{17}
$$

where $\boldsymbol{\alpha} = [\mathfrak{q}; \mathfrak{p}]$, $\boldsymbol{\beta} = [\mathfrak{y}; \mathfrak{z}]$, $\boldsymbol{\gamma} = [\widehat{\mathfrak{y}}; \widehat{\mathfrak{z}}; \mathfrak{u}]$, $\boldsymbol{b} = [b_1; -b_2]$, and $\{B_i\}_{i=1}^8$ are the corresponding blocks of the coefficient matrix in (16).

Due to the discontinuous nature of the approximation spaces $\boldsymbol{V_h}$ and W_h, the first two equations of (17) can be used to eliminate both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in an element-by-element fashion. As a consequence, we can write system (17) as

$$
\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}
=
\begin{bmatrix} G_1 & H_1 \\ G_2 & H_2 \end{bmatrix}
\begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{b} \end{bmatrix} \tag{18}
$$

and

$$
B_6\boldsymbol{\alpha} + B_7\boldsymbol{\beta} + B_8\boldsymbol{\gamma} = 0. \tag{19}
$$

We provide details on the element-by-element construction of G_1, G_2 and H_1, H_2 in the appendix. Next, we eliminate both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to obtain a reduced globally coupled equation for $\boldsymbol{\gamma}$ only:

$$
\mathbb{K}\boldsymbol{\gamma} = -\,\mathbb{F}\boldsymbol{b}, \tag{20}
$$

where

$$
\mathbb{K} = B_6 G_1 + B_7 G_2 + B_8 \quad \text{and} \quad \mathbb{F} = B_6 H_1 + B_7 H_2.
$$

Once $\boldsymbol{\gamma}$ is available, both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be recovered from (18).

Remark 3.1 For HDG methods, the standard approach is to first compute the local solver independently on each element and then assemble the global system. The process we follow here is to first assemble the global system and then reduce its dimension by simple block-diagonal algebraic operations. The two approaches are equivalent.

The introduction of the approximate boundary trace unknowns, the “hybridization” process, allows this reduction of the global system to take place; see [14] for more information.

Equation (18) says we can express the approximate scalar state variable and corresponding fluxes in terms of the approximate traces on the element boundaries. The global Eq. (20) only involves the approximate traces. Therefore, the high number of globally coupled degrees of freedom in the HDG method is significantly reduced. This is one excellent feature of HDG methods.

3.3 Discretize-then-optimize and optimize-then-discretize

To approximate the solution of the optimal control problem (1)–(3), in Sects. 3.1, 3.2 we first derived the first-order necessary optimality conditions at the PDE level and then used the HDG method to discretize the optimality system. This strategy is called the optimize-then-discretize (OD) approach.

Another strategy is the discretize-then-optimize (DO) approach. Here, one first discretizes the PDE optimization problem to obtain a finite dimensional optimization problem. An advantage of this approach is that existing optimization algorithms can be utilized. However, this approach can yield spurious numerical results if the DO and OD approaches do not yield equivalent results; see, e.g., [37,39].

Since the DO approach allows the use of existing optimization algorithms, it is important to devise algorithms for which discretization and optimization *commute*, i.e., the DO and OD approaches yield equivalent results. In this section, we prove the two HDG approaches are equivalent. We follow the technique used in [61], where we prove optimization and discretization commute for an embedded discontinuous Galerkin method for a convection diffusion distributed control problem.

For the DO approach, we consider the following HDG discrete cost functional

$$\min_{u_h \in M_h(\partial)} \frac{1}{2}\|y_h - y_d\|^2_{\mathcal{T}_h} + \frac{\gamma}{2}\|u_h\|^2_{\varepsilon_h^\partial}, \qquad \gamma > 0,$$

subject to the HDG discretized state equations for $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o) \in \boldsymbol{V}_h \times W_h \times M_h(o)$ given by

$$\begin{aligned}
(\boldsymbol{q}_h, \boldsymbol{r}_1)_{\mathcal{T}_h} - (y_h, \nabla\cdot\boldsymbol{r}_1)_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle u_h, \boldsymbol{r}_1\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} &= 0,\\
(\nabla\cdot\boldsymbol{q}_h, w_1)_{\mathcal{T}_h} + \langle h^{-1}P_M y_h, w_1\rangle_{\partial\mathcal{T}_h} - \langle h^{-1}\widehat{y}_h^o, w_1\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}&\\
-\langle h^{-1}u_h, w_1\rangle_{\varepsilon_h^\partial} &= (f, w_1)_{\mathcal{T}_h},\\
\langle \boldsymbol{q}_h\cdot\boldsymbol{n} + h^{-1}(P_M y_h - \widehat{y}_h^o), \mu_1\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} &= 0,
\end{aligned}$$

which hold for any $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{V}_h \times W_h \times M_h(o)$.

The Lagrangian functional defined for the adjoint variables $(\boldsymbol{p}_h, z_h, \widehat{z}_h^o) \in \boldsymbol{V}_h \times W_h \times M_h(o)$ is given by

$$
\begin{aligned}
\mathcal{L}_h(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{p}_h, z_h, \widehat{z}_h^o) &= \frac{1}{2}\|y_h - y_d\|_{\mathcal{T}_h}^2 + \frac{\gamma}{2}\|u_h\|_{\varepsilon_h^\partial}^2 \\
&+ (\boldsymbol{q}_h, \boldsymbol{p}_h)_{\mathcal{T}_h} - (y_h, \nabla\cdot\boldsymbol{p}_h)_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{p}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle u_h, \boldsymbol{p}_h\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} \\
&- (\nabla\cdot\boldsymbol{q}_h, z_h)_{\mathcal{T}_h} - \langle h^{-1}P_M y_h, z_h\rangle_{\partial\mathcal{T}_h} + \langle h^{-1}\widehat{y}_h^o, z_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}u_h, z_h\rangle_{\varepsilon_h^\partial} \\
&+ (f, z_h)_{\mathcal{T}_h} + \langle \boldsymbol{q}_h\cdot\boldsymbol{n} + h^{-1}(P_M y_h - \widehat{y}_h^o), \widehat{z}_h^o\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}. \qquad (21)
\end{aligned}
$$

Due to the linear-quadratic structure of the optimization problem, we obtain the optimality system by setting the partial Fréchet-derivatives of (21) with respect to the flux $\boldsymbol{q}_h$, state y_h, numerical trace $\widehat{y}_h^o$ and control u_h equal to zero. This gives the adjoint equations and optimality condition

$$
\begin{aligned}
\frac{\partial\mathcal{L}_h}{\partial\boldsymbol{q}_h}\boldsymbol{r}_2 &= (\boldsymbol{p}_h, \boldsymbol{r}_2)_{\mathcal{T}_h} - (z_h, \nabla\cdot\boldsymbol{r}_2)_{\mathcal{T}_h} + \langle \widehat{z}_h^o, \boldsymbol{r}_2\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0, \\
\frac{\partial\mathcal{L}_h}{\partial y_h}w_2 &= -(\nabla\cdot\boldsymbol{p}_h, w_2)_{\mathcal{T}_h} - \langle h^{-1}P_M z_h, w_2\rangle_{\partial\mathcal{T}_h} + \langle h^{-1}\widehat{z}_h^o, w_2\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} \\
&\quad + (y_h - y_d, w_2)_{\mathcal{T}_h} = 0, \\
\frac{\partial\mathcal{L}_h}{\partial\widehat{y}_h^o}\mu_2 &= \langle \boldsymbol{p}_h\cdot\boldsymbol{n} + h^{-1}(z_h - \widehat{z}_h^o), \mu_2\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0, \\
\frac{\partial\mathcal{L}_h}{\partial u_h}\mu_3 &= \langle \gamma u_h + \boldsymbol{p}_h\cdot\boldsymbol{n} + h^{-1}z_h, \mu_3\rangle_{\varepsilon_h^\partial} = 0,
\end{aligned}
$$

which hold for all $(\boldsymbol{r}_2, w_2, \mu_2, \mu_3) \in \boldsymbol{V}_h\times W_h\times M_h(o)\times M_h(\partial)$. Comparing the HDG discretized optimality system (14) with the above equations shows the two approaches are equivalent, i.e., OD = DO.

4 Error analysis

Next, we provide a convergence analysis of the above HDG method for the Dirichlet boundary control problem. Throughout this section, we assume $\Omega\subset\mathbb{R}^d$, $d = 2, 3$, is a bounded convex polyhedral domain, the regularity condition (8), (9) is satisfied, and h is bounded above by some positive constant. For the 2D case, recall Sect. 2 provides conditions on Ω and y_d guaranteeing the required regularity.

Through this section, we use $A \lesssim B$ to indicate that there exists a positive constant C such that $A \leq CB$, where C only depends on the polynomial degree k, the domain Ω and the shape regularity of the mesh.

4.1 Main result

First, we present the following main theoretical result of this work. Recall we assume the fractional Sobolev regularity exponents satisfy

$$
r_u > 1, \quad r_y > 1, \quad r_z > 2, \quad r_{\boldsymbol{q}} > 1/2, \quad r_{\boldsymbol{p}} > 1.
$$

Theorem 4.1 *For*

$$s_y = \min\{r_y, k+2\},\ s_z = \min\{r_z, k+2\},\ s_{\boldsymbol{q}} = \min\{r_{\boldsymbol{q}}, k+1\},\ s_{\boldsymbol{p}} = \min\{r_{\boldsymbol{p}}, k+1\},$$

we have

$$\begin{aligned}
\|u - u_h\|_{\varepsilon_h^\partial} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}\,,\\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}\,,\\
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-1} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-2} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-1} \|y\|_{s_y,\Omega}\,,\\
\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}\,,\\
\|z - z_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}\,.
\end{aligned}$$

Using the regularity results for the 2D case presented in Sect. 2, we obtain the following result.

Corollary 4.2 *Suppose $d = 2$ and $f = 0$ Let $\omega \in [\pi/3, 2\pi/3)$ be the largest interior angle of Γ, and define p_Ω, r_Ω by*

$$p_\Omega = \frac{2}{2 - \pi/\max\{\omega, \pi/2\}} \in (4, \infty], \quad r_\Omega = 1 + \frac{\pi}{\omega} \in (5/2, 4].$$

Assume $y_d \in L^p(\Omega) \cap H^{r-2}(\Omega)$ for all $p < p_\Omega$ and $r < r_\Omega$. If $k = 1$, then for any r satisfying $5/2 < r < \min\{3, r_\Omega\}$ we have

$$\begin{aligned}
\|u - u_h\|_{\varepsilon_h^\partial} &\lesssim h^{r-\frac{3}{2}} (\|\boldsymbol{p}\|_{H^{r-1}(\Omega)} + \|z\|_{H^r(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{r-\frac{3}{2}} (\|\boldsymbol{p}\|_{H^{r-1}(\Omega)} + \|z\|_{H^r(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} &\lesssim h^{r-2} (\|\boldsymbol{p}\|_{H^{r-1}(\Omega)} + \|z\|_{H^r(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h} &\lesssim h^{r-\frac{3}{2}} (\|\boldsymbol{p}\|_{H^{r-1}(\Omega)} + \|z\|_{H^r(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|z - z_h\|_{\mathcal{T}_h} &\lesssim h^{r-\frac{3}{2}} (\|\boldsymbol{p}\|_{H^{r-1}(\Omega)} + \|z\|_{H^r(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}).
\end{aligned}$$

Furthermore, if $k = 0$, then for any r as above we have

$$\begin{aligned}
\|u - u_h\|_{\varepsilon_h^\partial} &\lesssim h^{1/2} (\|\boldsymbol{p}\|_{H^1(\Omega)} + \|z\|_{H^2(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{1/2} (\|\boldsymbol{p}\|_{H^1(\Omega)} + \|z\|_{H^2(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h} &\lesssim h^{1/2} (\|\boldsymbol{p}\|_{H^1(\Omega)} + \|z\|_{H^2(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}),\\
\|z - z_h\|_{\mathcal{T}_h} &\lesssim h^{1/2} (\|\boldsymbol{p}\|_{H^1(\Omega)} + \|z\|_{H^2(\Omega)} + \|\boldsymbol{q}\|_{H^{r-2}(\Omega)} + \|y\|_{H^{r-1}(\Omega)}).
\end{aligned}$$

Note that $\min\{3, r_\Omega\}$ is always greater than 5/2, which guarantees a superlinear convergence rate for all variables except $\boldsymbol{q}$ if $k = 1$. Also, if Ω is a rectangle (i.e.,

$\omega = \pi/2$) and $y_d \in H^1(\Omega) \cap L^\infty(\Omega)$, then $r_\Omega = 3$ and we obtain an $O(h^{3/2-\varepsilon})$ convergence rate for u, y, z, and $\boldsymbol{p}$, and an $O(h^{1-\varepsilon})$ convergence rate for $\boldsymbol{q}$ for any $\varepsilon > 0$. For $k = 1$, the convergence rates are optimal for the control u and the flux $\boldsymbol{q}$, but suboptimal for the other variables. For $k = 0$, the convergence rates are suboptimal for all variables.

4.2 Preliminary material

Before we prove the main result, we discuss L^2 projections, an HDG operator $\mathcal{B}$, and the well-posedness of the HDG equations.

For any element K and boundary face e, we define the standard L^2 projections $\boldsymbol{\Pi} : [L^2(K)]^d \to [\mathcal{P}^k(K)]^d$, $\Pi : L^2(K) \to \mathcal{P}^{k+1}(K)$, and $P_M : L^2(e) \to \mathcal{P}^k(e)$, which satisfy

$$\begin{aligned} (\boldsymbol{\Pi q}, \boldsymbol{r})_K &= (\boldsymbol{q}, \boldsymbol{r})_K, && \forall \boldsymbol{r} \in [\mathcal{P}^k(K)]^d, \\ (\Pi u, w)_K &= (u, w)_K, && \forall w \in \mathcal{P}^{k+1}(K), \\ \langle P_M m, \mu\rangle_e &= \langle m, \mu\rangle_e, && \forall \mu \in \mathcal{P}^k(e). \end{aligned} \tag{22}$$

Note that $\boldsymbol{\Pi} : [L^2(\Omega)]^d \to \boldsymbol{V}_h$, $\Pi : L^2(\Omega) \to W_h$, and $P_M : L^2(\varepsilon_h) \to M_h$.
In the analysis, we use the following classical results:

$$\|\boldsymbol{q} - \boldsymbol{\Pi q}\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{q}}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega}, \qquad \|y - \Pi y\|_{\mathcal{T}_h} \lesssim h^{s_y} \|y\|_{s_y,\Omega}, \tag{23a}$$

$$\|y - \Pi y\|_{\partial\mathcal{T}_h} \lesssim h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}, \qquad \|\boldsymbol{q} - \boldsymbol{\Pi q}\|_{\partial\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{q}}-\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega}, \tag{23b}$$

$$\|w\|_{\partial\mathcal{T}_h} \lesssim h^{-\frac{1}{2}} \|w\|_{\mathcal{T}_h} \qquad \forall w \in W_h, \tag{23c}$$

where $s_{\boldsymbol{q}}$ and s_y are defined in Theorem 4.1. We note that (23a) follows directly from [54, Theorem 2.6], and the inverse inequality (23c) can be found in [60]. Since $y \in H^1(\Omega)$, the inequality for y in (23b) follows from an approximation result [54, Theorem 2.6], a trace inequality [4, Theorem 1.6.6], and the stability of the L^2 projection in the H^1 norm [22, Lemma 1.131]. The same proof works for the inequality for $\boldsymbol{q}$ in (23b) if $\boldsymbol{q}$ is also in $H^1(\Omega)$. However, if $q \in H^{r_{\boldsymbol{q}}}$ for some $r_{\boldsymbol{q}} \in (1/2, 1)$, then the inequality follows from an approximation result [27, Lemma A.5], a trace inequality [23, Lemma 7.2], and the stability of the L^2 projection in the $H^{r_{\boldsymbol{q}}}$ norm (which can be shown using an interpolation argument). We have the same projection error bounds for $\boldsymbol{p}$, z, and other variables.

To shorten lengthy equations, for any $(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_1, w_1, \mu_1) \in [\boldsymbol{V}_h \times W_h \times M_h(o)]^2$, we define the HDG operator $\mathcal{B}$ as follows:

$$\begin{aligned} &\mathcal{B}(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_1, w_1, \mu_1) \\ &\quad = (\boldsymbol{q}_h, \boldsymbol{r}_1)_{\mathcal{T}_h} - (y_h, \nabla\cdot \boldsymbol{r}_1)_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{r}_1 \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} \\ &\qquad - (\boldsymbol{q}_h, \nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1} P_M y_h, w_1\rangle_{\partial\mathcal{T}_h} \\ &\qquad - \langle h^{-1}\widehat{y}_h^o, w_1\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1}(P_M y_h - \widehat{y}_h^o), \mu_1\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}. \end{aligned} \tag{24}$$

By the definition of $\mathcal{B}$, we can rewrite the HDG formulation of the optimality system (13) as follows: find $(\boldsymbol{q}_h, \boldsymbol{p}_h, y_h, z_h, \widehat{y}_h^o, \widehat{z}_h^o, u_h) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o) \times M_h(\partial)$ such that

$$\mathcal{B}(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{r}_1, w_1, \mu_1) = -\langle u_h, \boldsymbol{r_1} \cdot \boldsymbol{n} - h^{-1} w_1\rangle_{\varepsilon_h^\partial} + (f, w_1)_{\mathcal{T}_h}, \tag{25a}$$

$$\mathcal{B}(\boldsymbol{p}_h, z_h, \widehat{z}_h^o; \boldsymbol{r}_2, w_2, \mu_2) = (y_h - y_d, w_2)_{\mathcal{T}_h}, \tag{25b}$$

$$\gamma^{-1}\langle \boldsymbol{p}_h \cdot \boldsymbol{n} + h^{-1} P_M z_h, \mu_3\rangle_{\varepsilon_h^\partial} = -\langle u_h, \mu_3\rangle_{\varepsilon_h^\partial} \tag{25c}$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2, w_1, w_2, \mu_1, \mu_2, \mu_3) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o) \times M_h(\partial)$.

Next, we present a basic property of the operator $\mathcal{B}$ and show the HDG Eq. (25) has a unique solution.

Lemma 4.3 *For any $(\boldsymbol{v}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h(o)$, we have*

$$\begin{aligned}\mathcal{B}(\boldsymbol{v}_h, w_h, \mu_h; \boldsymbol{v}_h, w_h, \mu_h) = (\boldsymbol{v}_h, \boldsymbol{v}_h)_{\mathcal{T}_h} + \langle h^{-1}(P_M w_h - \mu_h), P_M w_h - \mu_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} \\ + \langle h^{-1} P_M w_h, P_M w_h\rangle_{\varepsilon_h^\partial}.\end{aligned}$$

Proof By the definition of $\mathcal{B}$ in (24) and integration by parts, we have

$$\begin{aligned}&\mathcal{B}(\boldsymbol{v}_h, w_h, \mu_h; \boldsymbol{v}_h, w_h, \mu_h)\\ &\quad= (\boldsymbol{v}_h, \boldsymbol{v}_h)_{\mathcal{T}_h} - (w_h, \nabla\cdot\boldsymbol{v}_h)_{\mathcal{T}_h} + \langle \mu_h, \boldsymbol{v}_h \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - (\boldsymbol{v}_h, \nabla w_h)_{\mathcal{T}_h}\\ &\qquad+ \langle \boldsymbol{v}_h \cdot \boldsymbol{n} + h^{-1} P_M w_h, w_h\rangle_{\partial\mathcal{T}_h} - \langle h^{-1}\mu_h, w_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\ &\qquad- \langle \boldsymbol{v}_h \cdot \boldsymbol{n} + h^{-1}(P_M w_h - \mu_h), \mu_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\ &\quad= (\boldsymbol{v}_h, \boldsymbol{v}_h)_{\mathcal{T}_h} + \langle h^{-1} P_M w_h, w_h\rangle_{\partial\mathcal{T}_h} - \langle h^{-1}\mu_h, w_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\ &\qquad- \langle h^{-1}(P_M w_h - \mu_h), \mu_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\ &\quad= (\boldsymbol{v}_h, \boldsymbol{v}_h)_{\mathcal{T}_h} + \langle h^{-1}(P_M w_h - \mu_h), P_M w_h - \mu_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1} P_M w_h, P_M w_h\rangle_{\varepsilon_h^\partial}.\end{aligned}$$

□

Proposition 4.4 *There exists a unique solution of the HDG Eq. (25).*

Proof Since the system (25) is finite dimensional, we only need to prove the uniqueness. Therefore, we assume $y_d = f = 0$ and we show the system (25) only has the trivial solution.

First, by the definition of $\mathcal{B}$, we have

$$\begin{aligned}&\mathcal{B}(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{p}_h, -z_h, -\widehat{z}_h^o) + \mathcal{B}(\boldsymbol{p}_h, z_h, \widehat{z}_h^o; -\boldsymbol{q}_h, y_h, \widehat{y}_h^o)\\ &\quad= (\boldsymbol{q}_h, \boldsymbol{p}_h)_{\mathcal{T}_h} - (y_h, \nabla\cdot\boldsymbol{p}_h)_{\mathcal{T}_h} + \langle \widehat{y}_h^o, \boldsymbol{p}_h \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + (\boldsymbol{q}_h, \nabla z_h)_{\mathcal{T}_h}\\ &\qquad- \langle \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1} P_M y_h, z_h\rangle_{\partial\mathcal{T}_h} + \langle h^{-1}\widehat{y}_h^o, z_h\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\ &\qquad+ \langle \boldsymbol{q}_h \cdot \boldsymbol{n} + h^{-1}(P_M y_h - \widehat{y}_h^o), \widehat{z}_h^o\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - (\boldsymbol{p}_h, \boldsymbol{q}_h)_{\mathcal{T}_h} + (z_h, \nabla\cdot\boldsymbol{q}_h)_{\mathcal{T}_h}\end{aligned}$$

$$- \langle \widehat{z}_h^o, \boldsymbol{q}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} - (\boldsymbol{p}_h, \nabla y_h)_{\mathcal{T}_h} + \langle \boldsymbol{p}_h \cdot \boldsymbol{n} + h^{-1} P_M z_h, y_h \rangle_{\partial \mathcal{T}_h}$$
$$- \langle h^{-1} \widehat{z}_h^o, y_h \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} - \langle \boldsymbol{p}_h \cdot \boldsymbol{n} + h^{-1}(P_M z_h - \widehat{z}_h^o), \widehat{y}_h^o \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}.$$

Integrating by parts and using the properties of P_M in (22) gives

$$\mathcal{B}(\boldsymbol{q}_h, y_h, \widehat{y}_h^o; \boldsymbol{p}_h, -z_h, -\widehat{z}_h^o) + \mathcal{B}(\boldsymbol{p}_h, z_h, \widehat{z}_h^o; -\boldsymbol{q}_h, y_h, \widehat{y}_h^o) = 0.$$

Next, take $(\boldsymbol{r}_1, w_1, \mu_1) = (\boldsymbol{p}_h, -z_h, -\widehat{z}_h^o)$, $(\boldsymbol{r}_2, w_2, \mu_2) = (-\boldsymbol{q}_h, y_h, \widehat{y}_h^o)$, and $\mu_3 = -\gamma u_h$ in the HDG Eqs. (25a)–(25c), respectively, and sum to obtain

$$(y_h, y_h)_{\mathcal{T}_h} + \gamma \, \|u_h\|^2_{\varepsilon_h^\partial} = 0.$$

This implies $y_h = 0$ and $u_h = 0$ since $\gamma > 0$.

Next, taking $(\boldsymbol{r}_1, w_1, \mu_1) = (\boldsymbol{q}_h, y_h, \widehat{y}_h^o)$ and $(\boldsymbol{r}_2, w_2, \mu_2) = (\boldsymbol{p}_h, z_h, \widehat{z}_h^o)$ in Lemma 4.3 gives $\boldsymbol{q}_h = \boldsymbol{p}_h = \boldsymbol{0}$, $\widehat{y}_h^o = 0$, $P_M z_h = 0$ on ε_h^∂ and

$$P_M z_h - \widehat{z}_h^o = 0 \quad \text{on } \partial \mathcal{T}_h \backslash \varepsilon_h^\partial. \tag{26}$$

Substituting (26) into (13c), and remembering again $P_M z_h = 0$ on ε_h^∂, we get

$$-(z_h, \nabla \cdot \boldsymbol{r_2})_{\mathcal{T}_h} + \langle P_M z_h, \boldsymbol{r_2} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0.$$

Use the property of P_M in (22), integrate by parts, and take $\boldsymbol{r}_2 = \nabla z_h$ to obtain

$$(\nabla z_h, \nabla z_h)_{\mathcal{T}_h} = 0.$$

Thus, z_h is constant on each $K \in \mathcal{T}_h$. Therefore, $P_M z_h$ is constant on ∂K for each $K \in \mathcal{T}_h$. Since $P_M z_h = 0$ on ε_h^∂ and $P_M z_h = \widehat{z}_h^o = 0$ on $\partial \mathcal{T}_h \backslash \varepsilon_h^\partial$, this gives $z_h = 0$ on $\partial \mathcal{T}_h$. Therefore $z_h = 0$ since z_h is constant on each $K \in \mathcal{T}_h$.

4.3 Proof of main result

To prove the main result, we follow a similar strategy taken by Gong and Yan [29], see also [12,40,43], and introduce an auxiliary problem with the approximate control u_h in (25a) replaced by a projection of the exact optimal control. We first bound the error between the solutions of the auxiliary problem and the mixed weak form (4a)–(4e) of the optimality system. Then we bound the error between the solutions of the auxiliary problem and the HDG problem (25). A simple application of the triangle inequality then gives a bound on the error between the solutions of the HDG problem and the mixed form of the optimality system.

The precise form of the auxiliary problem is given as follows: find $(\boldsymbol{q}_h(u), \ \boldsymbol{p}_h(u), y_h(u), z_h(u), \widehat{y}_h^o(u), \widehat{z}_h^o(u)) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o)$ such that

$$\mathcal{B}(\boldsymbol{q}_h(u), y_h(u), \widehat{y}_h^o(u); \boldsymbol{r}_1, w_1, \mu_1) = -\langle P_M u, \boldsymbol{r_1} \cdot \boldsymbol{n} - h^{-1} w_1 \rangle_{\varepsilon_h^\partial} + (f, w_1)_{\mathcal{T}_h}, \tag{27a}$$

$$\mathcal{B}(\boldsymbol{p}_h(u), z_h(u), \widehat{z}_h^o(u); \boldsymbol{r}_2, w_2, \mu_2) = (y_h(u) - y_d, w_2)_{\mathcal{T}_h} \tag{27b}$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2, w_1, w_2, \mu_1, \mu_2) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o)$.

We split the proof of the main result, Theorem 4.1, in 7 steps. We begin by bounding the error between the solutions of the auxiliary problem and the mixed form (4a)–(4e) of the optimality system. We split the errors in the variables using the L^2 projections. In steps 1–3, we focus on the primary variables, i.e., the state y and the flux $\boldsymbol{q}$, and we use the following notation:

$$\begin{aligned}
\delta^{\boldsymbol{q}} &= \boldsymbol{q} - \boldsymbol{\Pi q}, & \varepsilon_h^{\boldsymbol{q}} &= \boldsymbol{\Pi q} - \boldsymbol{q}_h(u),\\
\delta^{y} &= y - \Pi y, & \varepsilon_h^{y} &= \Pi y - y_h(u),\\
\delta^{\widehat{y}} &= y - P_M y, & \varepsilon_h^{\widehat{y}} &= P_M y - \widehat{y}_h^o(u),\\
\widehat{\boldsymbol{\delta}}_1 &= \delta^{\boldsymbol{q}} \cdot \boldsymbol{n} + h^{-1} P_M \delta^y, & \widehat{\boldsymbol{\varepsilon}}_1 &= \varepsilon_h^{\boldsymbol{q}} \cdot \boldsymbol{n} + h^{-1}(P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}}).
\end{aligned} \tag{28}$$

4.3.1 Step 1: The error equation for part 1 of the auxiliary problem (27a)

Lemma 4.5 *For any $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{V}_h \times W_h \times M_h(o)$, we have*

$$\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^{y}, \varepsilon_h^{\widehat{y}}; \boldsymbol{r}_1, w_1, \mu_1) = -\langle \widehat{\boldsymbol{\delta}}_1, w_1 \rangle_{\partial \mathcal{T}_h} + \langle \widehat{\boldsymbol{\delta}}_1, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}. \tag{29}$$

Proof By the definition of the operator $\mathcal{B}$ in (24), we have

$$\begin{aligned}
&\mathcal{B}(\boldsymbol{\Pi q}, \Pi y, P_M y; \boldsymbol{r}_1, w_1, \mu_1)\\
&\quad= (\boldsymbol{\Pi q}, \boldsymbol{r_1})_{\mathcal{T}_h} - (\Pi y, \nabla \cdot \boldsymbol{r_1})_{\mathcal{T}_h} + \langle P_M y, \boldsymbol{r_1} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}\\
&\quad\quad- (\boldsymbol{\Pi q}, \nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{\Pi q} \cdot \boldsymbol{n} + h^{-1} P_M \Pi y, w_1 \rangle_{\partial \mathcal{T}_h}\\
&\quad\quad- \langle h^{-1} P_M y, w_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} - \langle \boldsymbol{\Pi q} \cdot \boldsymbol{n} - h^{-1} P_M \delta^y, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}.
\end{aligned}$$

By properties of the L^2 projections (22), we have

$$\begin{aligned}
\mathcal{B}(\boldsymbol{\Pi q}, \Pi y, P_M y; \boldsymbol{r}_1, w_1, \mu_1) &= (\boldsymbol{q}, \boldsymbol{r_1})_{\mathcal{T}_h} - (y, \nabla \cdot \boldsymbol{r_1})_{\mathcal{T}_h} + \langle y, \boldsymbol{r_1} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}\\
&\quad- (\boldsymbol{q}, \nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{q} \cdot \boldsymbol{n}, w_1 \rangle_{\partial \mathcal{T}_h} - \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, w_1 \rangle_{\partial \mathcal{T}_h}\\
&\quad+ \langle h^{-1} P_M \Pi y, w_1 \rangle_{\partial \mathcal{T}_h} - \langle h^{-1} P_M y, w_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}\\
&\quad- \langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}\\
&\quad+ \langle h^{-1} P_M \delta^y, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial}.
\end{aligned}$$

Note that the exact state y and exact flux $\boldsymbol{q}$ satisfy

$$\begin{aligned}
(\boldsymbol{q}, \boldsymbol{r}_1)_{\mathcal{T}_h} - (y, \nabla \cdot \boldsymbol{r}_1)_{\mathcal{T}_h} + \langle y, \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} &= -\langle u, \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\varepsilon_h^\partial},\\
-(\boldsymbol{q}, \nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{q} \cdot \boldsymbol{n}, w_1 \rangle_{\partial \mathcal{T}_h} &= (f, w_1)_{\mathcal{T}_h},\\
\langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} &= 0
\end{aligned}$$

for all $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{V}_h \times W_h \times M_h(o)$ and $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{V}_h \times W_h \times M_h$. Here, the last equation holds since $\boldsymbol{q} \in H(\text{div}, \Omega)$ and μ_1 is single-valued on each interior face. Then we have

$$\begin{aligned}\mathcal{B}(\boldsymbol{\Pi q}, \Pi y, P_M y; \boldsymbol{r}_1, w_1, \mu_1) = &- \langle u, \boldsymbol{r}_1 \cdot \boldsymbol{n}\rangle_{\varepsilon_h^\partial} + (f, w_1)_{\mathcal{T}_h} - \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, w_1\rangle_{\partial\mathcal{T}_h} \\ &+ \langle h^{-1} P_M \Pi y, w_1\rangle_{\partial\mathcal{T}_h} - \langle h^{-1} P_M y, w_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\ &+ \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, \mu_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle h^{-1} P_M \delta^y, \mu_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}.\end{aligned}$$

Subtract part 1 of the auxiliary problem (27a) from the above equality, use $y = u$ on ε_h^∂, and use $\langle P_M u, \boldsymbol{r}_1 \cdot \boldsymbol{n}\rangle_{\varepsilon_h^\partial} = \langle u, \boldsymbol{r}_1 \cdot \boldsymbol{n}\rangle_{\varepsilon_h^\partial}$ to obtain the result:

$$\begin{aligned}\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; \boldsymbol{r}_1, w_1, \mu_1) = &- \langle P_M u, h^{-1} w_1\rangle_{\varepsilon_h^\partial} - \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, w_1\rangle_{\partial\mathcal{T}_h} \\ &+ \langle h^{-1} P_M \Pi y, w_1\rangle_{\partial\mathcal{T}_h} - \langle h^{-1} P_M y, w_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\ &+ \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, \mu_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} + \langle h^{-1} P_M \delta^y, \mu_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\ = &- \langle \widehat{\boldsymbol{\delta}}_1, w_1\rangle_{\partial\mathcal{T}_h} + \langle \widehat{\boldsymbol{\delta}}_1, \mu_1\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}.\end{aligned}$$

□

4.3.2 Step 2: Estimate for $\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}$

We first provide a key inequality which was proven in Lemma 3.2 in [51]. In order to obtain the estimate for $\varepsilon_h^{\boldsymbol{q}}$, we first provide a key inequality which was proven in [51].

Lemma 4.6 *We have*

$$\begin{aligned}&\|\nabla \varepsilon_h^y\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|\varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|\varepsilon_h^y\|_{\varepsilon_h^\partial} \\ &\quad\lesssim \|\varepsilon_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^y\|_{\varepsilon_h^\partial}.\end{aligned}$$

Lemma 4.7 *We have*

$$\begin{aligned}&\|\varepsilon_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + h^{-1} \|P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}^2 + h^{-1} \|P_M \varepsilon_h^y\|_{\varepsilon_h^\partial}^2 \\ &\quad\lesssim h^{2s_{\boldsymbol{q}}} \|\boldsymbol{q}\|_{s^{\boldsymbol{q}},\Omega}^2 + h^{2s_y - 2} \|y\|_{s^y,\Omega}^2.\end{aligned} \tag{30}$$

Proof First, the basic property of $\mathcal{B}$ in Lemma 4.3 gives

$$\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; \varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}) = (\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^{\boldsymbol{q}})_{\mathcal{T}_h} + h^{-1} \|P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}^2 + h^{-1} \|P_M \varepsilon_h^y\|_{\varepsilon_h^\partial}^2.$$

Then, taking $(\boldsymbol{r}_1, w_1, \mu_1) = (\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}})$ in (29) in Lemma 4.5 gives

$$\begin{aligned}&(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^{\boldsymbol{q}})_{\mathcal{T}_h} + h^{-1} \|P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}^2 + h^{-1} \|P_M \varepsilon_h^y\|_{\varepsilon_h^\partial}^2 \\ &\quad= - \langle \widehat{\boldsymbol{\delta}}_1, \varepsilon_h^y - \varepsilon_h^{\widehat{y}}\rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} - \langle \widehat{\boldsymbol{\delta}}_1, \varepsilon_h^y\rangle_{\varepsilon_h^\partial}\end{aligned}$$

$$
\begin{aligned}
&= -\langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} - h^{-1} \langle \delta^y, P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \rangle_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\
&\quad - \langle \delta^{\boldsymbol{q}} \cdot \boldsymbol{n}, \varepsilon_h^y \rangle_{\varepsilon_h^\partial} - h^{-1} \langle \delta^y, P_M \varepsilon_h^y \rangle_{\varepsilon_h^\partial} \\
&\leq \left\| \delta^{\boldsymbol{q}} \right\|_{\partial\mathcal{T}_h} \| \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-1} \| \delta^y \|_{\partial\mathcal{T}_h} \| P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\
&\quad + \left\| \delta^{\boldsymbol{q}} \right\|_{\partial\mathcal{T}_h} \| \varepsilon_h^y \|_{\varepsilon_h^\partial} + h^{-1} \| \delta^y \|_{\partial\mathcal{T}_h} \| P_M \varepsilon_h^y \|_{\varepsilon_h^\partial} \\
&\leq h^{1/2} \left\| \delta^{\boldsymbol{q}} \right\|_{\partial\mathcal{T}_h} h^{-1/2} \| \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\
&\quad + h^{-1/2} \| \delta^y \|_{\partial\mathcal{T}_h} h^{-1/2} \| P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial} \\
&\quad + h^{1/2} \left\| \delta^{\boldsymbol{q}} \right\|_{\partial\mathcal{T}_h} h^{-1/2} \| \varepsilon_h^y \|_{\varepsilon_h^\partial} + h^{-1/2} \| \delta^y \|_{\partial\mathcal{T}_h} h^{-1/2} \| P_M \varepsilon_h^y \|_{\varepsilon_h^\partial}.
\end{aligned}
$$

By Young's inequality, Lemma 4.6, and the approximation properties of the L^2 projections in (23) we obtain

$$
\begin{aligned}
&\| \varepsilon_h^{\boldsymbol{q}} \|_{\mathcal{T}_h}^2 + h^{-1} \| P_M \varepsilon_h^y - \varepsilon_h^{\widehat{y}} \|_{\partial\mathcal{T}_h \backslash \varepsilon_h^\partial}^2 + h^{-1} \| P_M \varepsilon_h^y \|_{\varepsilon_h^\partial}^2 \\
&\qquad \lesssim h \left\| \delta^{\boldsymbol{q}} \right\|_{\partial\mathcal{T}_h}^2 + h^{-1} \left\| \delta^y \right\|_{\partial\mathcal{T}_h}^2 \\
&\qquad \lesssim h^{2s_{\boldsymbol{q}}} \left\| \boldsymbol{q} \right\|_{s^{\boldsymbol{q}},\Omega}^2 + h^{2s_y - 2} \left\| y \right\|_{s^y,\Omega}^2 . \qquad \square
\end{aligned}
$$

4.3.3 Step 3: Estimate for ε_h^y by a duality argument

Next, we introduce the dual problem for any given Θ in $L^2(\Omega)$:

$$
\begin{aligned}
\boldsymbol{\Phi} + \nabla \Psi &= 0 && \text{in } \Omega, \\
\nabla \cdot \boldsymbol{\Phi} &= \Theta && \text{in } \Omega, \\
\Psi &= 0 && \text{on } \Gamma.
\end{aligned} \tag{31}
$$

Since the domain Ω is convex, we have the following regularity estimate

$$
\| \boldsymbol{\Phi} \|_{H^1(\Omega)} + \| \Psi \|_{H^2(\Omega)} \lesssim \| \Theta \|_{\mathcal{T}_h} . \tag{32}
$$

Before we estimate ε_h^y we introduce the following notation, which is similar to the earlier notation in (28):

$$
\delta^{\boldsymbol{\Phi}} = \boldsymbol{\Phi} - \boldsymbol{\Pi} \boldsymbol{\Phi}, \quad \delta^\Psi = \Psi - \Pi \Psi, \quad \delta^{\widehat{\Psi}} = \Psi - P_M \Psi. \tag{33}
$$

By the regularity estimate (32) and the approximation properties of L^2 projections, we have the following bounds:

$$
\| \delta^{\boldsymbol{\Phi}} \|_{\partial\mathcal{T}_h} \lesssim h^{\frac{1}{2}} \| \Theta \|_{\mathcal{T}_h}, \quad \| \delta^\Psi \|_{\partial\mathcal{T}_h} \lesssim h^{\frac{3}{2}} \| \Theta \|_{\mathcal{T}_h}. \tag{34}
$$

Lemma 4.8 *We have*

$$
\| \varepsilon_h^y \|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{q}}+1} \left\| \boldsymbol{q} \right\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y} \left\| y \right\|_{s_y,\Omega} . \tag{35}
$$

Proof Consider the dual problem (31) and let $\Theta = \varepsilon_h^y$. In the definition (24) of $\mathcal{B}$, take $(\boldsymbol{r}_1, w_1, \mu_1)$ to be $(-\boldsymbol{\Pi\Phi}, \Pi\Psi, P_M\Psi)$ and use $\Psi = 0$ on ε_h^∂ to obtain

$$\begin{aligned}
&\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; -\boldsymbol{\Pi\Phi}, \Pi\Psi, P_M\Psi)\\
&= -(\varepsilon_h^{\boldsymbol{q}}, \boldsymbol{\Pi\Phi})_{\mathcal{T}_h} + (\varepsilon_h^y, \nabla\cdot\boldsymbol{\Pi\Phi})_{\mathcal{T}_h} - \langle \varepsilon_h^{\widehat{y}}, \boldsymbol{\Pi\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad - (\varepsilon_h^{\boldsymbol{q}}, \nabla\Pi\Psi)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{\varepsilon}}_1, \Pi\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle \widehat{\boldsymbol{\varepsilon}}_1, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad + \langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n} + h^{-1}P_M\varepsilon_h^y, \Pi\Psi\rangle_{\varepsilon_h^\partial}.
\end{aligned}\tag{36}$$

Next, integrating by parts, using properties of L^2 projections, and using $\nabla\cdot\boldsymbol{\Phi} = \varepsilon_h^y$ gives

$$\begin{aligned}
(\varepsilon_h^y, \nabla\cdot\boldsymbol{\Pi\Phi})_{\mathcal{T}_h} &= \langle \varepsilon_h^y, \boldsymbol{\Pi\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\nabla\varepsilon_h^y, \boldsymbol{\Pi\Phi})_{\mathcal{T}_h}\\
&= \langle \varepsilon_h^y, \boldsymbol{\Pi\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\nabla\varepsilon_h^y, \boldsymbol{\Phi})_{\mathcal{T}_h}\\
&= -\langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + (\varepsilon_h^y, \nabla\cdot\boldsymbol{\Phi})_{\mathcal{T}_h}\\
&= -\langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \|\varepsilon_h^y\|_{\mathcal{T}_h}^2.
\end{aligned}$$

Similarly, using $\boldsymbol{\Phi} + \nabla\Psi = \boldsymbol{0}$ gives

$$\begin{aligned}
-(\varepsilon_h^{\boldsymbol{q}}, \nabla\Pi\Psi)_{\mathcal{T}_h} &= -\langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \Pi\Psi\rangle_{\partial\mathcal{T}_h} + (\nabla\cdot\varepsilon_h^{\boldsymbol{q}}, \Pi\Psi)_{\mathcal{T}_h}\\
&= -\langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \Pi\Psi\rangle_{\partial\mathcal{T}_h} + (\nabla\cdot\varepsilon_h^{\boldsymbol{q}}, \Psi)_{\mathcal{T}_h}\\
&= -\langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \Pi\Psi\rangle_{\partial\mathcal{T}_h} + \langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h} - (\varepsilon_h^{\boldsymbol{q}}, \nabla\Psi)_{\mathcal{T}_h}\\
&= \langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, (P_M\Psi - \Pi\Psi)\rangle_{\partial\mathcal{T}_h} + (\varepsilon_h^{\boldsymbol{q}}, \boldsymbol{\Phi})_{\mathcal{T}_h}.
\end{aligned}$$

Then Eq. (36) becomes

$$\begin{aligned}
&\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; -\boldsymbol{\Pi\Phi}, \Pi\Psi, P_M\Psi)\\
&\quad = -(\varepsilon_h^{\boldsymbol{q}}, \boldsymbol{\Phi})_{\mathcal{T}_h} - \langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \|\varepsilon_h^y\|_{\mathcal{T}_h}^2 - \langle \varepsilon_h^{\widehat{y}}, \boldsymbol{\Pi\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad + \langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, P_M\Psi - \Pi\Psi\rangle_{\partial\mathcal{T}_h} + (\varepsilon_h^{\boldsymbol{q}}, \boldsymbol{\Phi})_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{\varepsilon}}_1, \Pi\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle \widehat{\boldsymbol{\varepsilon}}_1, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad + \langle h^{-1}P_M\varepsilon_h^y, \Pi\Psi\rangle_{\varepsilon_h^\partial}.
\end{aligned}$$

Since $\boldsymbol{\Phi}$ is single-valued on element faces, we have $\langle \varepsilon_h^{\widehat{y}}, \boldsymbol{\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0$. This gives $-\langle \varepsilon_h^{\widehat{y}}, \boldsymbol{\Pi\Phi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = \langle \varepsilon_h^{\widehat{y}}, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}$. Also, since $\langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h} = \langle \varepsilon_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h}$ and $\langle \widehat{\boldsymbol{\varepsilon}}_1, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = \langle \widehat{\boldsymbol{\varepsilon}}_1, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}$, we have

$$\begin{aligned}
&\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; -\boldsymbol{\Pi\Phi}, \Pi\Psi, P_M\Psi)\\
&\quad = -\langle \varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \|\varepsilon_h^y\|_{\mathcal{T}_h}^2 - h^{-1}\langle P_M\varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad - \langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} - h^{-1}\langle P_M\varepsilon_h^y, \delta^\Psi\rangle_{\varepsilon_h^\partial}.
\end{aligned}$$

On the other hand, Eq. (29) in Lemma 4.5 gives

$$\mathcal{B}(\varepsilon_h^{\boldsymbol{q}}, \varepsilon_h^y, \varepsilon_h^{\widehat{y}}; -\boldsymbol{\Pi\Phi}, \Pi\Psi, P_M\Psi) = -\langle \widehat{\boldsymbol{\delta}}_1, \Pi\Psi\rangle_{\partial\mathcal{T}_h} + \langle \widehat{\boldsymbol{\delta}}_1, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}.$$

Moreover,

$$\begin{aligned}
&\langle \widehat{\boldsymbol{\delta}}_1, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= \langle \delta^{\boldsymbol{q}}\cdot\boldsymbol{n} + h^{-1}P_M\delta^y, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= \langle \boldsymbol{q}\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle \boldsymbol{\Pi q}\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}P_M\delta^y, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= -\langle \boldsymbol{\Pi q}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}P_M\delta^y, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= \langle \boldsymbol{q}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle \boldsymbol{\Pi q}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}P_M\delta^y, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= \langle \widehat{\boldsymbol{\delta}}_1, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad= \langle \widehat{\boldsymbol{\delta}}_1, \Psi\rangle_{\partial\mathcal{T}_h},
\end{aligned}$$

where we used $\langle \boldsymbol{q}\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0$ and $\langle \boldsymbol{q}\cdot\boldsymbol{n}, \Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} = 0$, which hold since $\boldsymbol{q} \in H(\text{div}, \Omega)$, and we also used $\Psi = 0$ on ε_h^∂.

Comparing the above two equalities gives

$$\begin{aligned}
\|\varepsilon_h^y\|_{\mathcal{T}_h}^2 &= \langle \varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\langle P_M\varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad+ \langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} + h^{-1}\langle P_M\varepsilon_h^y, \delta^\Psi\rangle_{\varepsilon_h^\partial} + \langle \widehat{\boldsymbol{\delta}}_1, \delta^\Psi\rangle_{\partial\mathcal{T}_h}\\
&= \langle \varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\langle P_M\varepsilon_h^y - \varepsilon_h^{\widehat{y}}, \delta^\Psi\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\quad+ \langle \varepsilon_h^y, \delta^{\boldsymbol{\Phi}}\cdot\boldsymbol{n}\rangle_{\varepsilon_h^\partial} + h^{-1}\langle P_M\varepsilon_h^y, \delta^\Psi\rangle_{\varepsilon_h^\partial} + \langle \delta^{\boldsymbol{q}}\cdot\boldsymbol{n} + h^{-1}P_M\delta^y, \delta^\Psi\rangle_{\partial\mathcal{T}_h}\\
&\lesssim h^{-\frac{1}{2}}\|\varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\cdot h^{\frac{1}{2}}\|\delta^{\boldsymbol{\Phi}}\|_{\partial\mathcal{T}_h}\\
&\quad+ h^{-\frac{1}{2}}\|P_M\varepsilon_h^y - \varepsilon_h^{\widehat{y}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\cdot h^{-\frac{1}{2}}\|\delta^\Psi\|_{\partial\mathcal{T}_h}\\
&\quad+ h^{-\frac{1}{2}}\|\varepsilon_h^y\|_{\varepsilon_h^\partial}\cdot h^{\frac{1}{2}}\|\delta^{\boldsymbol{\Phi}}\|_{\partial\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^y\|_{\varepsilon_h^\partial}\cdot h^{-\frac{1}{2}}\|\delta^\Psi\|_{\partial\mathcal{T}_h}\\
&\quad+ \|\delta^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h}\cdot\|\delta^\Psi\|_{\partial\mathcal{T}_h} + h^{-1}\|\delta^y\|_{\partial\mathcal{T}_h}\cdot\|\delta^\Psi\|_{\partial\mathcal{T}_h}\\
&\lesssim (h^{s_{\boldsymbol{q}}+1}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\|y\|_{s_y,\Omega})\|\varepsilon_h^y\|_{\mathcal{T}_h},
\end{aligned}$$

where we used Lemmas 4.6, 4.7, and the bounds (34) to obtain the final inequality. □

As a consequence of Lemmas 4.7, 4.8, a simple application of the triangle inequality gives optimal convergence rates for $\|\boldsymbol{q} - \boldsymbol{q}_h(u)\|_{\mathcal{T}_h}$ and $\|y - y_h(u)\|_{\mathcal{T}_h}$:

Lemma 4.9

$$\|\boldsymbol{q} - \boldsymbol{q}_h(u)\|_{\mathcal{T}_h} \le \|\delta^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \|\varepsilon_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{q}}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-1}\|y\|_{s_y,\Omega}, \tag{37a}$$

$$\|y - y_h(u)\|_{\mathcal{T}_h} \le \|\delta^y\|_{\mathcal{T}_h} + \|\varepsilon_h^y\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{q}}+1}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\|y\|_{s_y,\Omega}. \tag{37b}$$

4.3.4 Step 4: The error equation for part 2 of the auxiliary problem (27b)

We continue to bound the error between the solutions of the auxiliary problem and the mixed form (4a)–(4e) of the optimality system. In steps 4–5, we focus on the dual variables, i.e., the state z and the flux $\boldsymbol{p}$. We split the errors in the variables using the L^2 projections, and we use the following notation.

$$\begin{aligned} \delta^{\boldsymbol{p}} &= \boldsymbol{p} - \boldsymbol{\Pi} \boldsymbol{p}, & \varepsilon_h^{\boldsymbol{p}} &= \boldsymbol{\Pi} \boldsymbol{p} - \boldsymbol{p}_h(u),\\ \delta^z &= z - \Pi z, & \varepsilon_h^z &= \Pi z - z_h(u),\\ \delta^{\widehat{z}} &= z - P_M z, & \varepsilon_h^{\widehat{z}} &= P_M z - \widehat{z}_h^o(u),\\ \widehat{\boldsymbol{\delta}}_2 &= \delta^{\boldsymbol{p}} \cdot \boldsymbol{n} + h^{-1} P_M \delta^z. \end{aligned} \tag{38}$$

The derivation of the error equation for part 2 of the auxiliary problem (27b) is similar to the analysis for part 1 of the auxiliary problem in step 1 in Sect. 4.3.1; the only difference is there is one more term $(y - y_h(u), w_2)_{\mathcal{T}_h}$ in the right hand side. Therefore, we state the result and omit the proof.

Lemma 4.10 *For any $(\boldsymbol{r}_2, w_2, \mu_2) \in \boldsymbol{V}_h \times W_h \times M_h(o)$, we have*

$$\begin{aligned} &\mathcal{B}(\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^z, \varepsilon_h^{\widehat{z}}, \boldsymbol{r}_2, w_2, \mu_2)\\ &\quad = -\langle \widehat{\boldsymbol{\delta}}_2, w_2 \rangle_{\partial \mathcal{T}_h} + \langle \widehat{\boldsymbol{\delta}}_2, \mu_2 \rangle_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + (y - y_h(u), w_2)_{\mathcal{T}_h}. \end{aligned} \tag{39}$$

4.3.5 Step 5: Estimate for $\boldsymbol{\varepsilon}_h^{\boldsymbol{p}}$ and $\boldsymbol{\varepsilon}_h^z$

Before we estimate $\varepsilon_h^{\boldsymbol{p}}$, we give the following discrete Poincaré inequality from [51, p. 354].

Lemma 4.11 *We have*

$$\|\varepsilon_h^z\|_{\mathcal{T}_h} \lesssim \|\nabla \varepsilon_h^z\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|\varepsilon_h^z\|_{\varepsilon_h^\partial}. \tag{40}$$

Lemma 4.12 *We have*

$$\begin{aligned} &\left\|\varepsilon_h^{\boldsymbol{p}}\right\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^z\|_{\varepsilon_h^\partial}\\ &\qquad \lesssim h^{s_{\boldsymbol{p}}} \|\boldsymbol{p}\|_{s^{\boldsymbol{p}},\Omega} + h^{s_z - 1} \|z\|_{s^z,\Omega} + h^{s_{\boldsymbol{q}}+1} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y} \|y\|_{s_y,\Omega},\\ &\|\varepsilon_h^z\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{p}}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z - 1} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+1} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y} \|y\|_{s_y,\Omega}. \end{aligned}$$

Proof First, we note the key inequality in Lemma 4.6 is valid with $(z, \boldsymbol{p}, \hat{z})$ in place of $(y, \boldsymbol{q}, \hat{y})$. This gives

$$\begin{aligned} &\|\nabla \varepsilon_h^z\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|\varepsilon_h^z\|_{\varepsilon_h^\partial}\\ &\quad \lesssim \|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial \mathcal{T}_h \backslash \varepsilon_h^\partial} + h^{-\frac{1}{2}} \|P_M \varepsilon_h^z\|_{\varepsilon_h^\partial}, \end{aligned} \tag{41}$$

which we use below. Next, the basic property of $\mathcal{B}$ in Lemma 4.3 gives

$$\mathcal{B}(\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^z, \varepsilon_h^{\widehat{z}}, \varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^z, \varepsilon_h^{\widehat{z}}) = (\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^{\boldsymbol{p}})_{\mathcal{T}_h} + h^{-1}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|^2_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\|P_M\varepsilon_h^z\|^2_{\varepsilon_h^\partial}.$$

Then taking $(\boldsymbol{r}_2, w_2, \mu_2) = (\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^z, \varepsilon_h^{\widehat{z}})$ in (39) in Lemmas 4.10 and 4.11 give

$$\begin{aligned}
&(\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^{\boldsymbol{p}})_{\mathcal{T}_h} + h^{-1}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|^2_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\|P_M\varepsilon_h^z\|^2_{\varepsilon_h^\partial}\\
&\quad= -\langle\widehat{\boldsymbol{\delta}}_2, \varepsilon_h^z - \varepsilon_h^{\widehat{z}}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - \langle\widehat{\boldsymbol{\delta}}_2, \varepsilon_h^z\rangle_{\varepsilon_h^\partial} + (y - y_h(u), \varepsilon_h^z)_{\mathcal{T}_h}\\
&\quad= -\langle\delta^{\boldsymbol{p}}\cdot\boldsymbol{n}, \varepsilon_h^z - \varepsilon_h^{\widehat{z}}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} - h^{-1}\langle\delta^z, P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad- \langle\delta^{\boldsymbol{p}}\cdot\boldsymbol{n}, \varepsilon_h^z\rangle_{\varepsilon_h^\partial} - h^{-1}\langle\delta^z, P_M\varepsilon_h^z\rangle_{\varepsilon_h^\partial} + (y - y_h(u), \varepsilon_h^z)_{\mathcal{T}_h}\\
&\quad\le \left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h}\|\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\|\delta^z\|_{\partial\mathcal{T}_h}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad+ \left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h}\|\varepsilon_h^z\|_{\varepsilon_h^\partial} + h^{-1}\|\delta^z\|_{\partial\mathcal{T}_h}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial}\\
&\qquad+ \|y - y_h(u)\|_{\mathcal{T}_h}\|\varepsilon_h^z\|_{\mathcal{T}_h}\\
&\quad\le h^{1/2}\left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h} h^{-1/2}\|\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\delta^z\|_{\partial\mathcal{T}_h}h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial}\\
&\qquad+ h^{1/2}\left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h} h^{-1/2}\|\varepsilon_h^z\|_{\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\delta^z\|_{\partial\mathcal{T}_h}h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial}\\
&\qquad+ \|y - y_h(u)\|_{\mathcal{T}_h}\|\varepsilon_h^z\|_{\mathcal{T}_h}\\
&\quad\lesssim h^{1/2}\left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h}(\|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial})\\
&\qquad+ h^{-\frac{1}{2}}\|\delta^z\|_{\partial\mathcal{T}_h}h^{-\frac{1}{2}}(\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial})\\
&\qquad+ \|y - y_h(u)\|_{\mathcal{T}_h}(\|\nabla\varepsilon_h^z\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\varepsilon_h^z\|_{\varepsilon_h^\partial})\\
&\quad\lesssim h^{1/2}\left\|\delta^{\boldsymbol{p}}\right\|_{\partial\mathcal{T}_h}(\|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial})\\
&\qquad+ h^{-\frac{1}{2}}\|\delta^z\|_{\partial\mathcal{T}_h}h^{-\frac{1}{2}}(\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial})\\
&\qquad+ \|y - y_h(u)\|_{\mathcal{T}_h}(\|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial}).
\end{aligned}$$

Applying Young's inequality and Lemma 4.9 gives

$$\begin{aligned}
&(\varepsilon_h^{\boldsymbol{p}}, \varepsilon_h^{\boldsymbol{p}})_{\mathcal{T}_h} + h^{-1}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|^2_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-1}\|P_M\varepsilon_h^z\|^2_{\varepsilon_h^\partial}\\
&\qquad\lesssim h\left\|\delta^{\boldsymbol{p}}\right\|^2_{\partial\mathcal{T}_h} + h^{-1}\|\delta^z\|^2_{\partial\mathcal{T}_h} + \|y_h(u) - y\|^2_{\mathcal{T}_h}\\
&\qquad\lesssim h^{2s_{\boldsymbol{p}}}\|\boldsymbol{p}\|^2_{s_{\boldsymbol{p}},\Omega} + h^{2s_z-2}\|z\|^2_{s_z,\Omega} + h^{2s_{\boldsymbol{q}}+2}\|\boldsymbol{q}\|^2_{s_{\boldsymbol{q}},\Omega} + h^{2s_y}\|y\|^2_{s_y,\Omega}.
\end{aligned}$$

This gives

$$\begin{aligned}
&\|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial}\\
&\qquad\lesssim h^{s_{\boldsymbol{p}}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-1}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+1}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\|y\|_{s_y,\Omega},
\end{aligned}$$

$$
\begin{aligned}
\|\varepsilon_h^z\|_{\mathcal{T}_h} &\lesssim \|\nabla\varepsilon_h^z\|_{\mathcal{T}_h} + +h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial} \\
&\lesssim \|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z - \varepsilon_h^{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\varepsilon_h^z\|_{\varepsilon_h^\partial} \\
&\lesssim h^{s_{\boldsymbol{p}}}\,\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-1}\,\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+1}\,\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\,\|y\|_{s_y,\Omega}\,.
\end{aligned}
$$

□

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

Lemma 4.13

$$
\begin{aligned}
\|\boldsymbol{p}-\boldsymbol{p}_h(u)\|_{\mathcal{T}_h} &\le \|\delta^{\boldsymbol{p}}\|_{\mathcal{T}_h} + \|\varepsilon_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} \\
&\lesssim h^{s_{\boldsymbol{p}}}\,\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-1}\,\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+1}\,\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\,\|y\|_{s_y,\Omega}\,,
\end{aligned} \tag{42a}
$$

$$
\begin{aligned}
\|z-z_h(u)\|_{\mathcal{T}_h} &\le \|\delta^{z}\|_{\mathcal{T}_h} + \|\varepsilon_h^{z}\|_{\mathcal{T}_h} \\
&\lesssim h^{s_{\boldsymbol{p}}}\,\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-1}\,\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+1}\,\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y}\,\|y\|_{s_y,\Omega}\,.
\end{aligned} \tag{42b}
$$

4.3.6 Step 6: Estimate for $\|u-u_h\|_{\varepsilon_h^\partial}$ and $\|y-y_h\|_{\mathcal{T}_h}$

Next, we bound the error between the solutions of the auxiliary problem and the HDG problem (25). We use these error bounds and the error bounds in Lemma 4.9, 4.12, 4.13 to obtain the main result.

For the remaining steps, we denote

$$
\begin{aligned}
\zeta_{\boldsymbol{q}} &= \boldsymbol{q}_h(u) - \boldsymbol{q}_h, \quad \zeta_y = y_h(u) - y_h, \quad \zeta_{\widehat{y}} = \widehat{y}_h^o(u) - \widehat{y}_h^o, \\
\zeta_{\boldsymbol{p}} &= \boldsymbol{p}_h(u) - \boldsymbol{p}_h, \quad \zeta_z = z_h(u) - z_h, \quad \zeta_{\widehat{z}} = \widehat{z}_h^o(u) - \widehat{z}_h^o.
\end{aligned}
$$

Subtracting the auxiliary problem and the HDG problem gives the following error equations

$$
\mathcal{B}(\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}; \boldsymbol{r}_1, w_1, \mu_1) = -\langle P_M u - u_h, \boldsymbol{r}_1\cdot\boldsymbol{n} - h^{-1}w_1\rangle_{\varepsilon_h^\partial} \tag{43a}
$$

$$
\mathcal{B}(\zeta_{\boldsymbol{p}}, \zeta_z, \zeta_{\widehat{z}}; \boldsymbol{r}_2, w_2, \mu_2) = (\zeta_y, w_2)_{\mathcal{T}_h} \tag{43b}
$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2, w_1, w_2, \mu_1, \mu_2) \in \boldsymbol{V}_h \times \boldsymbol{V}_h \times W_h \times W_h \times M_h(o) \times M_h(o)$.

Lemma 4.14 *We have*

$$
\begin{aligned}
&\|u-u_h\|^2_{\varepsilon_h^\partial} + \gamma^{-1}\,\|\zeta_y\|^2_{\mathcal{T}_h} \\
&\quad = \langle u + \gamma^{-1}\boldsymbol{p}_h(u)\cdot\boldsymbol{n} + \gamma^{-1}h^{-1}P_M z_h(u), u-u_h\rangle_{\varepsilon_h^\partial} \\
&\qquad - \langle u_h + \gamma^{-1}\boldsymbol{p}_h\cdot\boldsymbol{n} + \gamma^{-1}h^{-1}P_M z_h, u-u_h\rangle_{\varepsilon_h^\partial}.
\end{aligned}
$$

Proof First, we have

$$
\begin{aligned}
&\langle u + \gamma^{-1} \boldsymbol{p}_h(u)\cdot \boldsymbol{n} + \gamma^{-1} h^{-1} P_M z_h(u), u - u_h\rangle_{\varepsilon_h^\partial} \\
&\quad - \langle u_h + \gamma^{-1} \boldsymbol{p}_h \cdot \boldsymbol{n} + \gamma^{-1} h^{-1} P_M z_h, u - u_h\rangle_{\varepsilon_h^\partial} \\
&= \|u - u_h\|_{\varepsilon_h^\partial}^2 + \gamma^{-1}\langle \zeta_{\boldsymbol{p}} \cdot \boldsymbol{n} + h^{-1} P_M \zeta_z, u - u_h\rangle_{\varepsilon_h^\partial}.
\end{aligned}
$$

As in the proof of Proposition 4.4, it can be shown that

$$
\mathcal{B}(\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}; \zeta_{\boldsymbol{p}}, -\zeta_z, -\zeta_{\widehat{z}}) + \mathcal{B}(\zeta_{\boldsymbol{p}}, \zeta_z, \zeta_{\widehat{z}}; -\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}) = 0.
$$

On the other hand, use the error equations in (43) to obtain

$$
\begin{aligned}
\mathcal{B}(\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}; \zeta_{\boldsymbol{p}}, -\zeta_z, -\zeta_{\widehat{z}}) &+ \mathcal{B}(\zeta_{\boldsymbol{p}}, \zeta_z, \zeta_{\widehat{z}}; -\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}) \\
&= (\zeta_y, \zeta_y)_{\mathcal{T}_h} - \langle P_M u - u_h, \zeta_{\boldsymbol{p}} \cdot \boldsymbol{n} + h^{-1} \zeta_z\rangle_{\varepsilon_h^\partial} \\
&= (\zeta_y, \zeta_y)_{\mathcal{T}_h} - \langle u - u_h, \zeta_{\boldsymbol{p}} \cdot \boldsymbol{n} + h^{-1} P_M \zeta_z\rangle_{\varepsilon_h^\partial}.
\end{aligned}
$$

Comparing the above two equalities gives

$$
(\zeta_y, \zeta_y)_{\mathcal{T}_h} = \langle u - u_h, \zeta_{\boldsymbol{p}} \cdot \boldsymbol{n} + h^{-1} P_M \zeta_z\rangle_{\varepsilon_h^\partial}.
$$

Theorem 4.15 *We have*

$$
\begin{aligned}
\|u - u_h\|_{\varepsilon_h^\partial} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}, \\
\|y - y_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}.
\end{aligned}
$$

Proof Since $u + \gamma^{-1}\boldsymbol{p}\cdot\boldsymbol{n} = 0$ on ε_h^∂ by (7e) and $u_h + \gamma^{-1}\boldsymbol{p}_h \cdot \boldsymbol{n} + \gamma^{-1}h^{-1}P_M z_h = 0$ on ε_h^∂ by (13g) and (13k) we have

$$
\begin{aligned}
\|u - u_h\|_{\varepsilon_h^\partial}^2 + \gamma^{-1}\left\|\zeta_y\right\|_{\mathcal{T}_h}^2 &= \langle u + \gamma^{-1} \boldsymbol{p}_h(u)\cdot \boldsymbol{n} + \gamma^{-1} h^{-1} P_M z_h(u), u - u_h\rangle_{\varepsilon_h^\partial} \\
&= \langle \gamma^{-1}(\boldsymbol{p}_h(u) - \boldsymbol{p})\cdot \boldsymbol{n} + \gamma^{-1} h^{-1} P_M z_h(u), u - u_h\rangle_{\varepsilon_h^\partial} \\
&\lesssim (\left\|\boldsymbol{p}_h(u) - \boldsymbol{p}\right\|_{\partial\mathcal{T}_h} + h^{-1}\|P_M z_h(u)\|_{\varepsilon_h^\partial})\, \|u - u_h\|_{\varepsilon_h^\partial}.
\end{aligned}
$$

Next, use an inverse inequality and Lemma 4.12 to get

$$
\begin{aligned}
\left\|\boldsymbol{p}_h(u) - \boldsymbol{p}\right\|_{\partial\mathcal{T}_h} &\le \left\|\boldsymbol{p}_h(u) - \boldsymbol{\Pi}\boldsymbol{p}\right\|_{\partial\mathcal{T}_h} + \|\boldsymbol{\Pi}\boldsymbol{p} - \boldsymbol{p}\|_{\partial\mathcal{T}_h} \\
&\lesssim h^{-\frac{1}{2}} \left\|\boldsymbol{p}_h(u) - \boldsymbol{\Pi}\boldsymbol{p}\right\|_{\mathcal{T}_h} + h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} \\
&\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}} \|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}} \|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}} \|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} \\
&\quad + h^{s_y-\frac{1}{2}} \|y\|_{s_y,\Omega}.
\end{aligned}
$$

Also, use $z=0$ on ε_h^∂, Lemma 4.12, and properties of the L^2 projection to obtain

$$
\begin{aligned}
h^{-1}\|P_M z_h(u)\|_{\varepsilon_h^\partial} &= h^{-1}\|P_M z_h(u) - P_M \Pi z + P_M \Pi z - P_M z\|_{\varepsilon_h^\partial}\\
&\le h^{-1}(\|P_M \varepsilon_h^z\|_{\varepsilon_h^\partial} + \|\Pi z - z\|_{\varepsilon_h^\partial})\\
&\le h^{-1}(\|P_M \varepsilon_h^z\|_{\varepsilon_h^\partial} + \|\Pi z - z\|_{\partial\mathcal{T}_h})\\
&\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega}\\
&\quad + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

This gives

$$
\|u-u_h\|_{\varepsilon_h^\partial} \lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
$$

Moreover, we have

$$
\|\zeta_y\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
$$

Since $y-y_h = y - y_h(u) + \zeta_y$, by the triangle inequality and Lemma 4.9 we obtain

$$
\|y-y_h\|_{\mathcal{T}_h} \lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
$$

□

Note that in the final estimate for $\|y-y_h\|_{\mathcal{T}_h}$ in the above proof, the lower convergence rate for $\|\zeta_y\|_{\mathcal{T}_h}$ dominates the higher convergence rate for $\|y-y_h(u)\|_{\mathcal{T}_h}$ from Lemma 4.9. In Step 7 below, the convergence rates for $\|\zeta_{\boldsymbol{p}}\|_{\mathcal{T}_h}$ and $\|\zeta_z\|_{\mathcal{T}_h}$ will also dominate the error bounds for the variables $\boldsymbol{p}$ and z.

4.3.7 Step 7: Estimates for $\|\boldsymbol{q}-\boldsymbol{q}_h\|_{\mathcal{T}_h}$, $\|\boldsymbol{p}-\boldsymbol{p}_h\|_{\mathcal{T}_h}$ and $\|z-z_h\|_{\mathcal{T}_h}$

Lemma 4.16 *We have*

$$
\begin{aligned}
\|\zeta_{\boldsymbol{q}}\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-1}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-2}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-1}\|y\|_{s_y,\Omega},\\
\|\zeta_{\boldsymbol{p}}\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega},\\
\|\zeta_z\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

Proof By Lemma 4.3 and the error Eq. (43a), we have

$$
\begin{aligned}
&\mathcal{B}(\zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}}; \zeta_{\boldsymbol{q}}, \zeta_y, \zeta_{\widehat{y}})\\
&\quad = (\zeta_{\boldsymbol{q}}, \zeta_{\boldsymbol{q}})_{\mathcal{T}_h} + \langle h^{-1}(P_M\zeta_y - \zeta_{\widehat{y}}), P_M\zeta_y - \zeta_{\widehat{y}}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}P_M\zeta_y, P_M\zeta_y\rangle_{\varepsilon_h^\partial}\\
&\quad = -\langle P_M u - u_h, \zeta_{\boldsymbol{q}}\cdot\boldsymbol{n} - h^{-1}\zeta_y\rangle_{\varepsilon_h^\partial} = -\langle u-u_h, \zeta_{\boldsymbol{q}}\cdot\boldsymbol{n} - h^{-1}P_M\zeta_y\rangle_{\varepsilon_h^\partial}
\end{aligned}
$$

$$
\begin{aligned}
&\lesssim \|u-u_h\|_{\varepsilon_h^\partial} \left(\left\|\boldsymbol{\zeta}_{\boldsymbol{q}}\right\|_{\varepsilon_h^\partial} + h^{-1}\left\|P_M\zeta_y\right\|_{\varepsilon_h^\partial}\right)\\
&\lesssim h^{-\frac{1}{2}} \|u-u_h\|_{\varepsilon_h^\partial} \left(\left\|\boldsymbol{\zeta}_{\boldsymbol{q}}\right\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\left\|P_M\zeta_y\right\|_{\varepsilon_h^\partial}\right),
\end{aligned}
$$

which gives

$$
\begin{aligned}
\left\|\zeta_{\boldsymbol{q}}\right\|_{\mathcal{T}_h} &\lesssim h^{-\frac{1}{2}} \|u-u_h\|_{\varepsilon_h^\partial}\\
&\lesssim h^{s_{\boldsymbol{p}}-1}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-2}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-1}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

Next, we estimate $\zeta_{\boldsymbol{p}}$. By Lemma 4.3, the error Eq. (43b), we have

$$
\begin{aligned}
&\mathcal{B}(\zeta_{\boldsymbol{p}}, \zeta_z, \zeta_{\widehat{z}}; \zeta_{\boldsymbol{p}}, \zeta_z, \zeta_{\widehat{z}})\\
&\quad= (\zeta_{\boldsymbol{p}}, \zeta_{\boldsymbol{p}})_{\mathcal{T}_h} + \langle h^{-1}(P_M\zeta_z - \zeta_{\widehat{z}}), P_M\zeta_z - \zeta_{\widehat{z}}\rangle_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + \langle h^{-1}P_M\zeta_z, P_M\zeta_z\rangle_{\varepsilon_h^\partial}\\
&\quad= (\zeta_y, \zeta_z)_{\mathcal{T}_h}\\
&\quad\le \left\|\zeta_y\right\|_{\mathcal{T}_h}\|\zeta_z\|_{\mathcal{T}_h}\\
&\quad\lesssim \left\|\zeta_y\right\|_{\mathcal{T}_h}\left(\|\nabla\zeta_z\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|\zeta_z - \zeta_{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\zeta_z\|_{\varepsilon_h^\partial}\right)\\
&\quad\lesssim \left\|\zeta_y\right\|_{\mathcal{T}_h}\left(\|\zeta_{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|\zeta_z - \zeta_{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\zeta_z\|_{\varepsilon_h^\partial}\right),
\end{aligned}
$$

where we used the discrete Poincaré inequality in Lemma 4.11 and also Lemma 4.6. This implies

$$
\begin{aligned}
&\left\|\zeta_{\boldsymbol{p}}\right\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\zeta_z - \zeta_{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\zeta_z\|_{\varepsilon_h^\partial}\\
&\qquad\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

The discrete Poincaré inequality in Lemma 4.11 also gives

$$
\begin{aligned}
\|\zeta_z\|_{\mathcal{T}_h} &\lesssim \|\nabla\zeta_z\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|\zeta_z - \zeta_{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|\zeta_z\|_{\varepsilon_h^\partial}\\
&\lesssim \|\zeta_{\boldsymbol{p}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}}\|P_M\zeta_z - \zeta_{\widehat{z}}\|_{\partial\mathcal{T}_h\backslash\varepsilon_h^\partial} + h^{-\frac{1}{2}}\|P_M\zeta_z\|_{\varepsilon_h^\partial}\\
&\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

□

The above lemma along with the triangle inequality, Lemmas 4.9,4.13 complete the proof of the main result:

Theorem 4.17 *We have*

$$
\begin{aligned}
\left\|\boldsymbol{q}-\boldsymbol{q}_h\right\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-1}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-2}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-1}\|y\|_{s_y,\Omega},\\
\left\|\boldsymbol{p}-\boldsymbol{p}_h\right\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega},\\
\|z-z_h\|_{\mathcal{T}_h} &\lesssim h^{s_{\boldsymbol{p}}-\frac{1}{2}}\|\boldsymbol{p}\|_{s_{\boldsymbol{p}},\Omega} + h^{s_z-\frac{3}{2}}\|z\|_{s_z,\Omega} + h^{s_{\boldsymbol{q}}+\frac{1}{2}}\|\boldsymbol{q}\|_{s_{\boldsymbol{q}},\Omega} + h^{s_y-\frac{1}{2}}\|y\|_{s_y,\Omega}.
\end{aligned}
$$

5 Numerical experiments

For our numerical experiments, we test problems similar to the examples considered in [29]; see also [6,42,48].

We begin with a 2D example on a square domain $\Omega = [0, 1/4] \times [0, 1/4] \subset \mathbb{R}^2$. The largest interior angle is $\omega = \pi/2$, and so $r_\Omega = 3$ and $p_\Omega = \infty$. The data is chosen as

$$f = 0, \quad y_d = (x^2 + y^2)^s \quad \text{and} \quad \gamma = 1,$$

where $s = 10^{-5}$. Then $y_d \in H^1(\Omega) \cap L^\infty(\Omega)$. For the case $k = 1$, i.e., quadratic polynomials are used for the scalar variables, and linear polynomials are used for the flux variables and the boundary trace variables, Corollary 4.2 in Sect. 4 gives the convergence rates

$$\begin{aligned} \|y - y_h\|_{0,\Omega} &= O(h^{3/2-\varepsilon}), \qquad \|z - z_h\|_{0,\Omega} = O(h^{3/2-\varepsilon}), \\ \|\boldsymbol{q} - \boldsymbol{q}_h\|_{0,\Omega} &= O(h^{1-\varepsilon}), \qquad \|\boldsymbol{p} - \boldsymbol{p}_h\|_{0,\Omega} = O(h^{3/2-\varepsilon}), \end{aligned}$$

and

$$\|u - u_h\|_{0,\Gamma} = O(h^{3/2-\varepsilon}).$$

For $k = 0$, i.e., linear polynomials are used for the scalar variables, and piecewise constant functions are used for the other variables, Corollary 4.2 gives convergence rates of $O(h^{1/2})$ for all variables.

Since we do not have an explicit expression for the exact solution, we solved the problem numerically for a triangulation with 262,144 elements, i.e., $h = 2^{-12}\sqrt{2}$ and compared this reference solution against other solutions computed on meshes with larger h.

The numerical results for $k = 1$ are shown in Table 1. The convergence rates observed for $\|\boldsymbol{q} - \boldsymbol{q}_h\|_{0,\Omega}$ and $\|u - u_h\|_{0,\Gamma}$ are in agreement with our theoretical results, while the convergence rates for $\|\boldsymbol{p} - \boldsymbol{p}_h\|_{0,\Omega}$, $\|y - y_h\|_{0,\Omega}$, and $\|z - z_h\|_{0,\Omega}$ are higher than our theoretical results. A similar phenomena can be observed in [29,48]. Only one work explained the above phenomena: May, Rannacher, and Vexler in [42] used a duality argument to obtain improved convergence rates for the state and dual state with the standard finite element method. To the best of our knowledge, it is not clear how to apply this technique to standard mixed finite element methods or the HDG method.

The numerical results for $k = 0$ are shown in Table 2. The convergence rates for all variables are higher than the $O(h^{1/2})$ rate from Corollary 4.2. As indicated in Sect. 4, the $O(h^{1/2})$ rate from Corollary 4.2 for $k = 0$ is suboptimal for all variables. Improving the numerical analysis for both the $k = 1$ and $k = 0$ cases is an interesting problem to be considered in the future.

For illustration, in Fig. 1 we plot the state y, adjoint state z, and their fluxes $\boldsymbol{q}$ and $\boldsymbol{p}$ computed using $k = 1$. The 2D regularity result in Sect. 2 indicates that the

Table 1 2D example, $k = 1$: error of control u, state y, adjoint state z, and their fluxes $\boldsymbol{q}$ and $\boldsymbol{p}$

$h/\sqrt{2}$	2^{-4}	$1/2^{-5}$	2^{-6}	2^{-7}	2^{-8}
$\\|\boldsymbol{q}-\boldsymbol{q}_h\\|_{0,\Omega}$	4.1343e-02	2.1025e−02	1.0677e−02	5.3865e−03	2.6959e−03
Order	–	0.9756	0.9776	0.9871	0.9986
$\\|\boldsymbol{p}-\boldsymbol{p}_h\\|_{0,\Omega}$	1.3463e−03	3.8638e−04	1.0849e−04	2.9862e−05	8.0969e−06
Order	–	1.8009	1.8325	1.8612	1.8828
$\\|y-y_h\\|_{0,\Omega}$	5.4609e−04	1.3647e−04	3.4763e−05	8.8037e−06	2.2236e−06
Order	–	2.0005	1.9730	1.9814	1.9852
$\\|z-z_h\\|_{0,\Omega}$	1.9671e−05	2.6887e−06	3.7026e−07	5.0372e−08	6.7767e−09
Order	–	2.8711	2.8603	2.8778	2.8940
$\\|u-u_h\\|_{0,\Gamma}$	7.3053e−03	2.6902e−03	9.7764e−04	3.5178e−04	1.2569e−04
Order	–	1.4412	1.4603	1.4746	1.4849

Table 2 2D example, $k = 0$: error of control u, state y, adjoint state z, and their fluxes $\boldsymbol{q}$ and $\boldsymbol{p}$

$h/\sqrt{2}$	2^{-4}	$1/2^{-5}$	2^{-6}	2^{-7}	2^{-8}
$\\|\boldsymbol{q}-\boldsymbol{q}_h\\|_{0,\Omega}$	4.7552e−02	3.4107e−02	2.1082e−02	1.2281e−02	6.9039e−03
Order	–	0.47942	0.69409	0.77961	0.83090
$\\|\boldsymbol{p}-\boldsymbol{p}_h\\|_{0,\Omega}$	1.6793e−03	9.8644e−04	5.2097e−04	2.6498e−04	1.3302e−04
Order	–	0.76759	0.92104	0.97531	0.99429
$\\|y-y_h\\|_{0,\Omega}$	7.3260e−04	3.2546e−04	1.0577e−04	3.1075e−05	8.7640e−06
Order	–	1.1706	1.6215	1.7671	1.8261
$\\|z-z_h\\|_{0,\Omega}$	7.3656e−05	2.0645e−05	5.4062e−06	1.3718e−06	3.4375e−07
Order	–	1.8350	1.9331	1.9785	1.9967
$\\|u-u_h\\|_{0,\Gamma}$	7.4915e−03	4.6700e−03	2.5730e−03	1.3539e−03	6.9528e−04
Order	–	0.68183	0.85996	0.92630	0.96148

primary flux $\boldsymbol{q}$ can have low regularity. In this example, it does indeed appear that $\boldsymbol{q}$ has singularities at the corners of the domain. We plot the computed control in Fig. 2. These figures can be compared to similar plots in [6,48].

Next, we consider a 3D extension of the 2D example above. The domain is a cube $\Omega = [0, 1/32] \times [0, 1/32] \times [0, 1/32]$, and the data is chosen as

$$f = 0, \quad y_d = (x^2 + y^2 + z^2)^s \quad \text{and} \quad \gamma = 1,$$

where $s = -1/4 + 10^{-5}$, so that $y_d \in H^1(\Omega)$. In this case, we did not attempt to determine the regularity of the control and other variables; we simply present the numerical results here.

As in the 2D example above, we do not have an explicit expression for the exact solution. Therefore, we solved the problem numerically using $k = 1$ for a triangulation with 196,608 tetrahedrons, i.e., $h = 2^{-12}\sqrt{3}$ and compared this reference solution against other solutions computed on meshes with larger h. The numerical results are

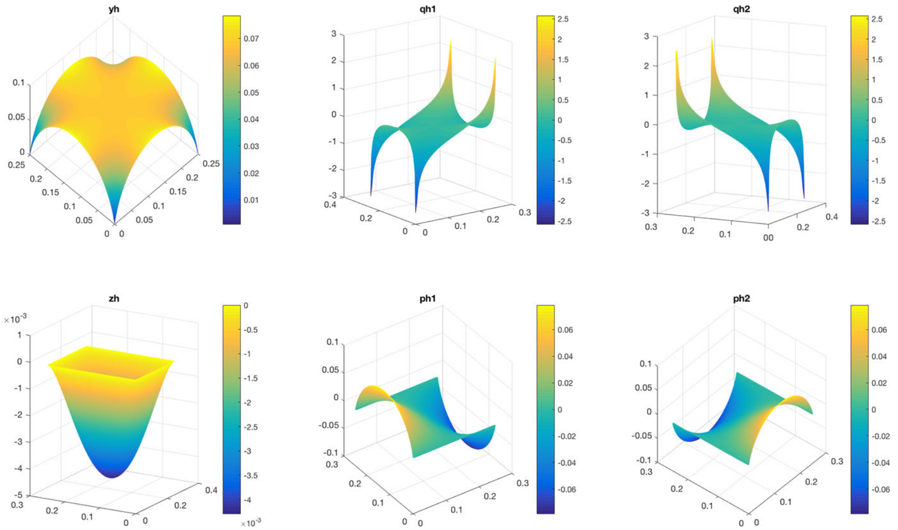

Fig. 1 2D example, $k = 1$: The computed primary state y_h, the primary flux $\boldsymbol{q}_h$, the dual state z_h, and the dual flux $\boldsymbol{p}_h$

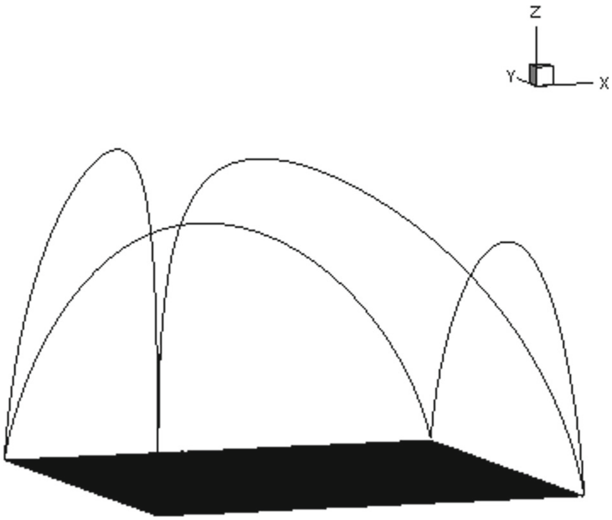

Fig. 2 The optimal control u_h for the 2D example

shown in Table 3. The observed convergence rates for all variables are similar to the results for the 2D example above.

6 Conclusions

We proposed an HDG method to approximate the solution of an optimal Dirichlet boundary control problem for the Poisson equation. We obtained an optimal superlin-

Table 3 3D example, $k = 1$: error of control u, state y, adjoint state z, and their fluxes $\boldsymbol{q}$ and $\boldsymbol{p}$

$h/\sqrt{3}$	2^{-6}	2^{-7}	2^{-8}	2^{-9}
$\Vert\boldsymbol{q}-\boldsymbol{q}_h\Vert_{0,\Omega}$	9.2640e−03	5.2580e−03	2.7462e−03	1.2475e−03
Order	–	0.81712	0.93706	1.1384
$\Vert\boldsymbol{p}-\boldsymbol{p}_h\Vert_{0,\Omega}$	3.5425e−05	1.2283e−05	3.8463e−06	1.1022e−06
Order	–	1.5281	1.6751	1.8032
$\Vert y-y_h\Vert_{0,\Omega}$	1.6040e−05	4.5070e−06	1.2191e−06	2.9781e−07
Order	–	1.8314	1.8864	2.0333
$\Vert z-z_h\Vert_{0,\Omega}$	7.8545e−08	1.3058e−08	2.0042e−09	2.8775e−10
Order	–	2.5886	2.7039	2.8001
$\Vert u-u_h\Vert_{0,\Gamma}$	4.5932e−04	1.8934e−04	7.1955e−05	2.4123e−05
Order	–	1.2785	1.3958	1.5767

ear rate of convergence for the control in 2D under certain assumptions on the domain, the target state y_d, and the polynomial degree of the HDG method. Numerical experiments confirmed our theoretical results for this superlinear convergence rate when piecewise linear polynomials were used to approximate the control (Table 1).

Our results indicate HDG methods have potential for solving more complex Dirichlet boundary control problems. We plan to investigate HDG methods for Dirichlet boundary control of other PDEs, including convection dominated diffusion problems and fluid flows. These problems may involve solutions with large gradients or shocks, and it is natural to consider HDG methods for such problems.

Acknowledgements The authors thank the referees for their comments, which helped improve the manuscript. W. Hu was supported in part by a postdoctoral fellowship for the annual program on Control Theory and its Applications at the Institute for Mathematics and its Applications (IMA) at the University of Minnesota, and also the DIG and FY 2018 ASR+1 Program at Oklahoma State University. J. Singler and Y. Zhang were supported in part by National Science Foundation grant DMS-1217122. J. Singler and Y. Zhang thank the IMA for funding research visits, during which some of this work was completed. X. Zheng thanks Missouri University of Science and Technology for hosting him as a visiting scholar; some of this work was completed during his research visit. The authors thank Bernardo Cockburn for many valuable conversations.

A Local solver

By simple algebraic operations in Eq. (17), we obtain the following formulas for the matrices G_1, G_2, H_1, and H_2 in (18):

$$
\begin{aligned}
G_1 &= B_1^{-1}B_2(B_4 + B_2^T B_1^{-1}B_2)^{-1}(B_5 + B_2^T B_1^{-1}B_3) - B_1^{-1}B_3,\\
G_2 &= -(B_4 + B_2^T B_1^{-1}B_2)^{-1}(B_5 + B_2^T B_1^{-1}B_3),\\
H_1 &= -B_1^{-1}B_2(B_4 + B_2^T B_1^{-1}B_2)^{-1},\\
H_2 &= (B_4 + B_2^T B_1^{-1}B_2)^{-1}.
\end{aligned}
$$

In general, this process is impractical; however, for the HDG method described in this work, these matrices can be easily computed. This is one of the advantages of the HDG method. We briefly describe this process below.

Since the spaces $\boldsymbol{V}_h$ and W_h consist of discontinuous polynomials, some of the system matrices are block diagonal and each block is small and symmetric positive definite. Let us call a matrix of this form a SSPD block diagonal matrix. The inverse of a SSPD block diagonal matrix is another SSPD block diagonal matrix, and the inverse can be easily constructed by computing the inverse of each small block. Furthermore, the inverse of each small block can be computed independently; and therefore computing the inverse can be easily done in parallel.

It can be checked that B_1 is a SSPD block diagonal matrix, and therefore B_1^{-1} is easily computed and is also a SSPD block diagonal matrix. Therefore, the matrices G_1, G_2, H_1, and H_2 are easily computed if $B_4 + B_2^T B_1^{-1} B_2$ is also easily inverted. We show below that this is the case.

First, it can be checked that B_2 is block diagonal with small blocks, but the blocks are not symmetric or definite. This implies $B_2^T B_1^{-1} B_2$ is block diagonal with small nonnegative definite blocks. Next, $B_4 = \begin{bmatrix} A_5 & 0 \\ -A_4 & A_5 \end{bmatrix}$, where A_4 and A_5 are both SSPD block diagonal. Due to the structure of B_1 and B_2, the matrix $B_2^T B_1^{-1} B_2 + B_4$ has the form $\begin{bmatrix} C_1 & 0 \\ -A_4 & C_2 \end{bmatrix}$, where C_1 and C_2 are SSPD block diagonal. The inverse can be easily computed using the formula

$$\begin{bmatrix} C_1 & 0 \\ -A_4 & C_2 \end{bmatrix}^{-1} = \begin{bmatrix} C_1^{-1} & 0 \\ C_2^{-1} A_4 C_1^{-1} & C_2^{-1} \end{bmatrix}.$$

Furthermore, C_1^{-1}, C_2^{-1} and $C_2^{-1} A_4 C_1^{-1}$ are both SSPD block diagonal.

References

1. 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)
2. Arada, N., Raymond, J.-P.: Dirichlet boundary control of semilinear parabolic equations. I. Problems with no state constraints. Appl. Math. Optim. **45**(2), 125–143 (2002)
3. Belgacem, F.B., El Fekih, H., Metoui, H.: Singular perturbation for the Dirichlet boundary control of elliptic problems. Math. Model. Numer. Anal. **37**(5), 850–883 (2003)
4. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, volume 15 of Texts in Applied Mathematics. Springer, New York (2008)
5. Casas, E., Mateos, M., Raymond, J.-P.: Penalization of Dirichlet optimal control problems. ESAIM Control Optim. Calc. Var. **15**(4), 782–809 (2009)
6. 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)
7. Cesmelioglu, A., Cockburn, B., Qiu, W.: Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier–Stokes equations. Math. Comput. **86**(306), 1643–1670 (2017)
8. Chamakuri, N., Engwer, C., Kunisch, K.: Boundary control of bidomain equations with state-dependent switching source functions in the ionic model. J. Comput. Phys. **273**, 227–242 (2014)

9. Chang, L., Gong, W., Yan, N.: Weak boundary penalization for Dirichlet boundary control problems governed by elliptic equations. J. Math. Anal. Appl. **453**(1), 529–557 (2017)
10. Chen, G. Pi, L., Zhang, Y.: A new ensemble HDG method for parameterized convection diffusion PDEs. 2019. arXiv:1910.10295
11. Chen, Y., Cockburn, B., Dong, B.: Superconvergent HDG methods for linear, stationary, third-order equations in one-space dimension. Math. Comput. **85**(302), 2715–2742 (2016)
12. Chen, Y.: Superconvergence of mixed finite element methods for optimal control problems. Math. Comput. **77**(263), 1269–1291 (2008)
13. Chowdhury, S., Gudi, T., Nandakumaran, A.K.: Error bounds for a Dirichlet boundary control problem based on energy spaces. Math. Comput. **86**(305), 1103–1126 (2017)
14. Cockburn, B.: Static condensation, hybridization, and the devising of the HDG methods. In: Building bridges: connections and challenges in modern approaches to numerical partial differential equations, volume 114 of Lect. Notes Comput. Sci. Eng., pp. 129–177. Springer, Cham (2016)
15. 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)
16. 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)
17. Cockburn, B., Mustapha, K.: A hybridizable discontinuous Galerkin method for fractional diffusion problems. Numer. Math. **130**(2), 293–314 (2015)
18. Cockburn, B., Shen, J.: A hybridizable discontinuous Galerkin method for the p-Laplacian. SIAM J. Sci. Comput. **38**(1), A545–A566 (2016)
19. Danilin, A.R.: Asymptotics of the solution of a problem of optimal boundary control of a flow through a part of the boundary. Tr. Inst. Mat. Mekh. **20**(4), 116–127 (2014)
20. de los Reyes, J.C., Kunisch, K.: A semi-smooth Newton method for control constrained boundary optimal control of the Navier–Stokes equations. Nonlinear Anal. **62**(7), 1289–1316 (2005)
21. 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)
22. Ern, A., Guermond, J.-L.: Theory and Practice of Finite Elements. Springer, New York (2004)
23. Ern, A., Guermond, J.-L.: Finite element quasi-interpolation and best approximation. ESAIM Math. Model. Numer. Anal. **51**(4), 1367–1385 (2017)
24. Fursikov, A.V., Gunzburger, M.D., Hou, L.S.: Boundary value problems and optimal boundary control for the Navier–Stokes system: the two-dimensional case. SIAM J. Control Optim. **36**(3), 852–894 (1998)
25. Fursikov, A.V., Gunzburger, M.D., Hou, L.S.: Optimal Dirichlet control and inhomogeneous boundary value problems for the unsteady Navier–Stokes equations. In: Proceedings of the Control and partial differential equations (Marseille-Luminy, 1997), volume 4 of ESAIM , pp. 97–116. Society for Industrial and Applied Mathematics, Paris (1998)
26. Fursikov, A.V., Gunzburger, M.D., Hou, L.S.: Optimal boundary control for the evolutionary Navier–Stokes system: the three-dimensional case. SIAM J. Control Optim. **43**(6), 2191–2232 (2005)
27. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier–Stokes Equations. Springer, Berlin (1986)
28. Gong, W., Li, B.: Improved error estimates for semi-discrete finite element solutions of parabolic Dirichlet boundary control problems. IMA J. Numer. Anal. To appear
29. 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)
30. 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)
31. Gunzburger, M.D., Manservisi, S.: The velocity tracking problem for Navier–Stokes flows with boundary control. SIAM J. Control Optim. **39**(2), 594–634 (2000)
32. Gunzburger, M.D., Hou, L.S., 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)

33. Gunzburger, M.D., Nicolaides, R.A.: An algorithm for the boundary control of the wave equation. Appl. Math. Lett. **2**(3), 225–228 (1989)
34. Hinze, M., Kunisch, K.: Second order methods for boundary control of the instationary Navier–Stokes system. ZAMM Z. Angew. Math. Mech. **84**(3), 171–187 (2004)
35. Hou, L.S., Ravindran, S.S.: A penalized Neumann control approach for solving an optimal Dirichlet control problem for the Navier–Stokes equations. SIAM J. Control Optim. **36**(5), 1795–1814 (1998)
36. Lehrenfeld, C.: Hybrid Discontinuous Galerkin Methods for Incompressible Flow Problems. Master's thesis, RWTH Aachen, May (2010)
37. Leykekhman, D.: Investigation of commutative properties of discontinuous Galerkin methods in pde constrained optimal control problems. J. Sci. Comput. **53**(3), 483–511 (2012)
38. Li, B., Xie, X.: Analysis of a family of HDG methods for second order elliptic problems. J. Comput. Appl. Math. **307**, 37–51 (2016)
39. Liu, J., Wang, Z.: Non-commutative discretize-then-optimize algorithms for elliptic PDE-constrained optimal control problems. J. Comput. Appl. Math. **362**, 596–613 (2019)
40. Liu, W., Yang, D., Yuan, L., Ma, C.: Finite element approximations of an optimal control problem with integral state constraint. SIAM J. Numer. Anal. **48**(3), 1163–1185 (2010)
41. Mateos, M., Neitzel, I.: Dirichlet control of elliptic state constrained problems. Comput. Optim. Appl. **63**(3), 825–853 (2016)
42. 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)
43. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control Optim. **43**(3), 970–985 (2004)
44. Mustapha, K., Nour, M., Cockburn, B.: Convergence and superconvergence analyses of HDG methods for time fractional diffusion problems. Adv. Comput. Math. **42**(2), 377–393 (2016)
45. Nguyen, N.C., Peraire, J., Cockburn, B.: An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations. J. Comput. Phys. **228**(9), 3232–3254 (2009)
46. Nguyen, N.C., Peraire, J., Cockburn, B.: An implicit high-order hybridizable discontinuous Galerkin method for nonlinear convection-diffusion equations. J. Comput. Phys. **228**(23), 8841–8855 (2009)
47. Nguyen, N.C., Peraire, J., Cockburn, B.: A hybridizable discontinuous Galerkin method for Stokes flow. Comput. Methods Appl. Mech. Eng. **199**(9–12), 582–597 (2010)
48. Of, G., Phan, T.X., Steinbach, O.: An energy space finite element approach for elliptic Dirichlet boundary control problems. Numer. Math. **129**(4), 723–748 (2015)
49. Oikawa, I.: A hybridized discontinuous Galerkin method with reduced stabilization. J. Sci. Comput. **65**(1), 327–340 (2015)
50. Qiu, W., Shen, J., Shi, K.: An HDG method for linear elasticity with strong symmetric stresses. Math. Comput. **87**(309), 69–93 (2018)
51. Qiu, W., Shi, K.: An HDG method for convection diffusion equation. J. Sci. Comput. **66**(1), 346–357 (2016)
52. Qiu, W., Shi, K.: A superconvergent HDG method for the incompressible Navier–Stokes equations on general polyhedral meshes. IMA J. Numer. Anal. **36**(4), 1943–1967 (2016)
53. Ravindran, S.S.: Finite element approximation of Dirichlet control using boundary penalty method for unsteady Navier–Stokes equations. ESAIM Math. Model. Numer. Anal. **51**(3), 825–849 (2017)
54. Rivière, B.: Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2008)
55. Sayas, F.-J., Brown, T.S., Hassell, M.E.: Variational Techniques for Elliptic Partial Differential Equations. CRC Press, Boca Raton, FL (2019)
56. Spasov, Y., Kunisch, K.: Dynamical system based optimal control of incompressible fluids. Boundary control. Eur. J. Mech. B Fluids **25**(2), 153–163 (2006)
57. Stanglmeier, M., Nguyen, N.C., Peraire, J., Cockburn, B.: An explicit hybridizable discontinuous Galerkin method for the acoustic wave equation. Comput. Methods Appl. Mech. Eng. **300**, 748–769 (2016)
58. Stavre, R.: A boundary control problem for the blood flow in venous insufficiency. The general case. Nonlinear Anal. Real World Appl. **29**, 98–116 (2016)
59. Vexler, B.: Finite element approximation of elliptic Dirichlet optimal control problems. Numer. Funct. Anal. Optim. **28**(7–8), 957–973 (2007)

60. Warburton, T., Hesthaven, J.S.: On the constants in hp-finite element trace inverse inequalities. Comput. Methods Appl. Mech. Eng. **192**(25), 2765–2773 (2003)
61. Zhang, X., Zhang, Y., Singler, J.R.: An optimal EDG method for distributed control of convection diffusion PDEs. Int. J. Numer. Anal. Model. **16**(4), 519–542 (2019)