IMA Journal of Numerical Analysis (2019) **00**, 1–24
doi:10.1093/imanum/drz003

On the error estimates of a hybridizable discontinuous Galerkin method for second-order elliptic problem with discontinuous coefficients

Gang Chen
School of Mathematics Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China
cglwdm@uestc.edu.cn

and

Jintao Cui*
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
*Corresponding author: jintao.cui@polyu.edu.hk

[Received on 17 January 2018; revised on 4 December 2018]

Hybridizable discontinuous Galerkin (HDG) methods retain the main advantages of standard discontinuous Galerkin (DG) methods, including their flexibility in meshing, ease of design and implementation, ease of use within an *hp*-adaptive strategy and preservation of local conservation of physical quantities. Moreover, HDG methods can significantly reduce the number of degrees of freedom, resulting in a substantial reduction of computational cost. In this paper, we study an HDG method for the second-order elliptic problem with discontinuous coefficients. The numerical scheme is proposed on general polygonal and polyhedral meshes with specially designed stabilization parameters. Robust *a priori* and *a posteriori* error estimates are derived without a full elliptic regularity assumption. The proposed *a posteriori* error estimators are proved to be efficient and reliable without a quasi-monotonicity assumption on the diffusion coefficient.

Keywords: hybridizable discontinuous Galerkin methods; *a priori* error estimates; *a posteriori* error estimates; discontinuous coefficient.

1. Introduction

Let $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3) be a polygonal or polyhedral domain with Lipschitz boundary $\partial\Omega := \Gamma = \overline{\Gamma_D} \cup \overline{\Gamma_N}$, where $meas(\Gamma_D) > 0$ and $\Gamma_D \cap \Gamma_N = \emptyset$. We consider the following second-order elliptic problem:

$$\begin{cases} a^{-1}\boldsymbol{p} - \nabla u = \mathbf{0}, & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{p} = f, & \text{in } \Omega, \\ u = g_D, & \text{on } \Gamma_D, \\ \boldsymbol{p} \cdot \boldsymbol{n} = g_N, & \text{on } \Gamma_N, \end{cases} \tag{1.1}$$

where $f \in L^2(\Omega)$, $g_D \in L^2(\Gamma_D)$ and $g_N \in L^2(\Gamma_N)$ are given scalar-valued functions; $\boldsymbol{n}$ is the outward unit normal vector; and the diffusion coefficient $a := a(\boldsymbol{x}) \in L^\infty(\Omega)$ is positive and piecewise constant on polygonal/polyhedral subdomains of Ω with possible large jumps across subdomain boundaries (interfaces).

For problem (1.1) with a discontinuous coefficient a or on a nonconvex domain Ω, the solutions may not be piecewise H^2 smooth. Especially when the coefficient a is discontinuous, the solutions of (1.1) may only have H^{1+s} regularity with small $s > 0$ (Kellogg, 1975; Grisvard, 1985), and are not piecewise H^{1+s} smooth with $s > 1/2$. However, the standard *a priori* error analysis of the discontinuous Galerkin (DG) methods requires the solution to be piecewise H^{1+s} smooth with $s > 1/2$. This theoretical gap is filled by the work of Cai *et al.* (2011), where the authors relax the regularity requirement to $s > 0$.

In the robust analysis of *a posteriori* error estimates, a quasi-monotonicity assumption on the diffusion coefficient is usually required. The concept of quasi-monotonicity has been introduced and exploited in Dryja *et al.* (1996) to obtain the robust interpolation properties of finite element methods in terms of weighted norms. Recently, robust *a posteriori* error estimates were given in Cai *et al.* (2017) without the quasi-monotonicity assumption. The error analysis only requires the solution to be in H^{1+s} with $s > 0$.

Since the late 1970s, DG methods have become increasingly popular due to their attractive features, including their flexibility in meshing, and preserving local conservation of physical quantities: see Arnold (1982) and Arnold *et al.* (2002) for elliptic boundary value problems. DG methods are also sui for parallel computation and ease of use within an *hp*-adaptive strategy. As pointed out in Demkowicz & Gopalakrishnan (2011), an inconvenient feature of DG methods is that they may require the penalization parameter to be 'sufficiently' large (practically unknown) for stability. This inconvenience was avoided by local discontinuous Galerkin (LDG) methods (Cockburn & Shu, 1998; Castillo *et al.*, 2000; Cockburn *et al.*, 2005; Carrero *et al.*, 2006), which have an additional property that fluxes can be eliminated locally. Later, hybridizable discontinuous Galerkin (HDG) methods (Cockburn *et al.*, 2009, 2010) were devised, which can also overcome this difficulty. HDG methods retain the advantages of standard DG methods and can significantly reduce the number of degrees of freedom, therefore allowing for a substantial reduction of computational cost. In Cockburn *et al.* (2010), an HDG method for second-order elliptic problems was studied, where the analysis requires H^2 regularity. In Li & Xie (2016), another HDG method for second-order elliptic problems was analyzed under H^1 regularity, where the constants in error estimates depend on the lower and upper bounds of the diffusion coefficient. The error analyses in both Cockburn *et al.* (2010) and Li & Xie (2016) are all derived based on simplicial meshes. HDG methods that allow polygonal meshes were first proposed in Lehrenfeld (2010) for elliptic problems. This approach has been extended and extensively studied for different problems, such as convection–diffusion problems (Qiu & Shi, 2016a), Navier–Stokes equations (Qiu & Shi, 2016b), Maxwell's equations (Chen *et al.*, 2017), elasticity problems (Qiu *et al.*, 2018), etc.

A posteriori error estimates for DG methods for (1.1) with H^1 regularity and smooth coefficient were given in Gudi (2010). The error analysis therein cannot be extended to the case of a non-smooth coefficient unless quasi-monotonicity of the coefficient is assumed. In Wihler & Rivière (2011), *a posteriori* error estimates for DG methods based on $W^{2,p}(\Omega), p \in (1,2]$ regularity for (1.1) with smooth coefficient were derived in two dimensions. In Di Pietro & Ern (2012), *a posteriori* error estimates for DG methods based on $W^{\frac{2d}{d+2},p}(\Omega), p \in (1,2]$ regularity for (1.1) with non-smooth coefficient were provided. Note that by Sobolev embedding theory, the regularity requirements in Wihler & Rivière (2011) and Di Pietro & Ern (2012) are actually stronger than $H^{1+s}(\Omega)$, $s > 0$. Recently, *a posteriori* error estimates for HDG methods for second-order elliptic problems with smooth coefficients were studied in Cockburn & Zhang (2012, 2013). The error analysis therein is also based on simplicial meshes.

In this paper, we propose and analyze an HDG method on general polygonal or polyhedral meshes for second-order elliptic problems with discontinuous coefficients. Robust *a priori* and *a posteriori* error estimates are given under low regularity assumptions. Our numerical scheme uses piecewise-polynomial approximations of degrees $k+1$, k and k ($k \geqslant 0$) for the scalar, the flux and the scalar on the inter-element boundaries, respectively. The *a posteriori* error estimates provide global upper bounds and local lower bounds for the error in terms of the error estimator, without the quasi-monotonicity assumption. Numerical experiments show that the errors converge at the same rates as for H^2-regular problems.

The rest of this paper is organized as follows: in section 2, we introduce our notational conventions and derive the *a priori* error estimates for the HDG method. In Section 3 we present the *a posteriori* error analysis for the HDG method. In Section 3, several numerical experiments are performed in order to confirm the theoretical results.

Throughout this paper, we use C to denote a positive constant independent of mesh size and a, which may take on different values at each occurrence. We use $a \lesssim b$ ($a \gtrsim b$) to represent $a \leqslant Cb$ ($a \geqslant Cb$), and $a \sim b$ to represent $a \lesssim b \lesssim a$.

2. Notation and the HDG method

2.1 *Notation*

For any bounded domain $\Lambda \subset \mathbb{R}^s$ ($s = d, d-1$), let $H^m(\Lambda)$ and $H_0^m(\Lambda)$ denote the usual Sobolev spaces on Λ, and $\|\cdot\|_{m,\Lambda}$ ($|\cdot|_{m,\Lambda}$, resp.) denote the norm (semi-norm, resp.) on these spaces. We use $(\cdot,\cdot)_{m,\Lambda}$ to denote the inner product of $H^m(\Lambda)$, with $(\cdot,\cdot)_\Lambda := (\cdot,\cdot)_{0,\Lambda}$. When $\Lambda = \Omega$, we denote $\|\cdot\|_m := \|\cdot\|_{m,\Omega}$, $|\cdot|_m := |\cdot|_{m,\Omega}$ and $(\cdot,\cdot) := (\cdot,\cdot)_\Omega$. In particular, when $\Lambda \in \mathbb{R}^{d-1}$, we use $\langle \cdot,\cdot\rangle_\Lambda$ to replace $(\cdot,\cdot)_\Lambda$. For an integer $k \geqslant 0$, $\mathbb{P}_k(\Lambda)$ denotes the set of all polynomials defined on Λ with degree less than or equal to k.

Let $\mathscr{T}_h = \bigcup\{T\}$ be a shape regular partition (to be defined later) of the domain Ω consisting of arbitrary polygons or polyhedra for $d = 2$ or 3, respectively. Note that $\mathscr{T}_h$ can be a conforming partition or a nonconforming partition, which allows hanging nodes.

For each $T \in \mathscr{T}_h$, we let h_T be the infimum of the diameters of circles (or spheres) containing T and denote the mesh size $h := \max_{T\in\mathscr{T}_h} h_T$. An edge (or face) E on the boundary ∂T of T is called a proper edge (or face) if all endpoints (or vertices) of the edge (or face) E are nodes of $\mathscr{T}_h$ and no other nodes of $\mathscr{T}_h$ are on E. In Fig. 1, for example EF, FH and HI are proper edges, while EH, FI and EI are not. Let $\mathscr{E}_h = \bigcup\{E\}$ be the union of all proper edges (faces) of $T \in \mathscr{T}_h$ and $\mathscr{E}_h^B = \mathscr{E}_h \cap \Gamma$ ($\mathscr{E}_h^D = \mathscr{E}_h \cap \Gamma_D$, $\mathscr{E}_h^N = \mathscr{E}_h \cap \Gamma_N$) be the union of all proper edges (faces) of $T \in \mathscr{T}_h$ on boundary Γ (Γ_D, Γ_N). Let $\mathscr{E}_h^I$ be the set of all interior proper edges (faces). We denote by h_E the length of edge E if $d = 2$ and the infimum of the diameters of circles containing face E if $d = 3$. For all $T \in \mathscr{T}_h$ and $E \in \mathscr{E}_h$, we denote by $\boldsymbol{n}_T$ and $\boldsymbol{n}_E$ the unit outward normal vectors along ∂T and E, respectively. Let $[\![v]\!]$ and $\{\!\{v\}\!\}$ denote the usual jump and mean values of a function v across every proper edge $E \in \mathscr{E}_h$.

The partition $\mathscr{T}_h$ is called shape regular when the following conditions hold:

- **M1** (Star-shaped elements). For each element $T \in \mathscr{T}_h$, there exists a positive constant θ_* and a point $M_T \in T$ such that T is star-shaped with respect to every point inside the circle (or sphere) whose center is M_T and radius is $\theta_* h_T$.
- **M2** (Edges or faces). For each element $T \in \mathscr{T}_h$, there exists a positive constant l_* such that the distance between any two vertices (including the hanging nodes) is greater than or equal to $l_* h_T$.

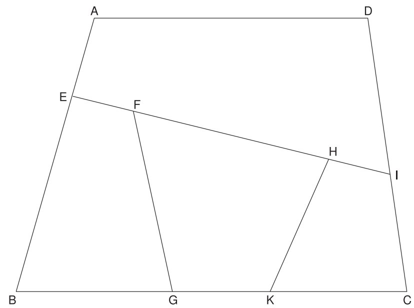

FIG. 1. Demonstration of edges in a two-dimensional mesh.

We also assume the partition $\mathscr{T}_h$ satisfies the following compatibility conditions:

- **T1**. Each boundary edge (or face) $E \in \mathscr{E}_h^B$ belongs to Γ_D or Γ_N.
- **T2**. For each element $T \in \mathscr{T}_h$, $a_T := a|_T$ is a constant.

When $d = 2$, for each $T \in \mathscr{T}_h$, we connect M_T and the vertices of T (including the hanging nodes) to get a set of triangles $w(T)$. When $d = 3$, for each face $E \subset \partial T$, we choose one vertex A of E and connect A and the other vertices of E to get a set of triangles $v(E)$; then we connect M_T and all vertices of the triangles in $v(E)$ to get a set of tetrahedrons $w(T)$. Let $\mathscr{M}_h := \bigcup_{T\in\mathscr{T}_h} w(T)$ and $\mathscr{F}_h$ be the union of all the edges (faces) of $\mathscr{M}_h$. Note that $\mathscr{M}_h$ is shape regular due to **M1** and **M2**.

We use ∇_h and $\nabla_h\cdot$ to denote the broken gradient and broken divergence with respect to $\mathscr{T}_h$ or $\mathscr{M}_h$. The following inverse inequality and trace inequality will be used in the error analysis.

LEMMA 2.1 For all $T \in \mathscr{T}_h$ and any given nonnegative integer j, the following inequalities hold true:

$$|w|_{1,T} \lesssim h_T^{-1}\|w\|_{0,T} \qquad \forall\, w \in \mathbb{P}_j(T), \tag{2.1}$$

$$\|w\|_{0,\partial T} \lesssim h_T^{-1/2}\|w\|_{0,T} + h_T^{1/2}|w|_{1,T} \qquad \forall\, w \in H^1(T). \tag{2.2}$$

Proof. By using an inverse inequality on the shape regular simplicial mesh $\mathscr{M}_h$, we have

$$|w|_{1,T}^2 = \sum_{M\in\mathscr{M}_h: M\subset T} |w|_{1,M}^2 \lesssim \sum_{M\in\mathscr{M}_h: M\subset T} h_M^{-1}\|w\|_{0,M}^2 \lesssim h_T^{-1}\|w\|_{0,T}^2 \qquad \forall\, w \in \mathbb{P}_j(T),$$

which proves (2.1).

By using a trace inequality on the shape regular simplicial mesh $\mathscr{M}_h$, for all $w \in H^1(T)$, it holds that

$$\begin{aligned}\|w\|_{0,\partial T}^2 &\leqslant \sum_{M\in\mathscr{M}_h: M\subset T} \|w\|_{0,\partial M}^2 \\ &\lesssim \sum_{M\in\mathscr{M}_h: M\subset T} \left(h_M^{-1}\|w\|_{0,M}^2 + h_M|w|_{1,M}^2\right) \\ &\lesssim h_T^{-1}\|w\|_{0,T}^2 + h_T|w|_{1,T}^2,\end{aligned}$$

which proves (2.2). □

Let $\gamma_T := \frac{h_T}{\rho_{T,\max}}$ be the chunkiness parameter of $T \in \mathcal{T}_h$, where $\rho_{T,\max}$ denotes the supremum of the radius of a sphere with respect to that T is star-shaped. Then, in view of **M1**, we have $2 \leqslant \gamma_T \leqslant \frac{h_T}{\theta_* h_T} = \theta_*^{-1}$, i.e. γ_T is independent of h_T. Thus, from Brenner & Scott (2008, Lemma 4.3.8) we obtain the following estimate.

LEMMA 2.2 For all $T \in \mathcal{T}_h$ and $v \in H^m(T)$ with $m \geqslant 1$, there exists $I_{m-1}v \in \mathbb{P}_{m-1}(T)$ such that

$$|v - I_{m-1}v|_{s,T} \lesssim h_T^{m-s}|v|_{m,T}, \quad \text{for } 0 \leqslant s \leqslant m. \tag{2.3}$$

In order to derive error estimates in Sections 3 and 4, we introduce the following results from Cai *et al.* (2017).

LEMMA 2.3 Let E be an edge (or face) of $T \in \mathcal{T}_h$, $\boldsymbol{n}_T$ be the unit vector normal to E and $s > 0$. Assume that v is a given function in $H^{1+s}(T)$ and $\Delta v \in L^2(T)$. Then for any $w_h \in \mathbb{P}_j(T)$ with a fixed nonnegative integer j, we have

$$\langle \nabla v \cdot \boldsymbol{n}_T, w_h \rangle_E \lesssim C h_E^{-1/2} \|w_h\|_{0,E} \left(\|\nabla v\|_{0,T} + h_T \|\Delta v\|_{0,T} \right).$$

REMARK 2.4 In Cai *et al.* (2017, Lemma 2.7), the above lemma holds for simplicial elements. Note that for every $T \in \mathcal{T}_h$, we can decompose T into several simplexes whose diameters are of order h_T; hence, Lemma 2.3 holds on T.

2.2 *An HDG finite element method*

For any $T \in \mathcal{T}_h$, $E \in \mathcal{E}_h$ and any nonnegative integer j, let $\Pi_j^o : L^2(T) \to \mathbb{P}_j(T)$ and $\Pi_j^\partial : L^2(E) \to \mathbb{P}_j(E)$ be the usual L^2-projection operators. Vector and tensor analogs of Π_j^o and Π_j^∂ are also denoted by Π_j^o and Π_j^∂, respectively.

For any integer $k \geqslant 0$, we introduce the following finite-dimensional spaces:

$$\begin{aligned}
V_h &:= \left\{ v_h \in L^2(\Omega) : v_h|_T \in \mathbb{P}_{k+1}(T),\ \forall\, T \in \mathcal{T}_h \right\}, \\
\widehat{V}_h &:= \left\{ \widehat{v}_h \in L^2(\mathcal{E}_h) : \widehat{v}_h|_E \in \mathbb{P}_k(E)],\ \forall\, E \in \mathcal{E}_h \right\}, \\
\widehat{V}_h^{\widetilde{g}} &:= \left\{ \widehat{v}_h \in \widehat{V}_h : \widehat{v}_h|_E = \Pi_k^\partial \widetilde{g},\ \forall\, E \in \mathcal{E}_h^D \right\}, \text{with } \widetilde{g} = 0,\ g_D, \\
\boldsymbol{Q}_h &:= \left\{ \boldsymbol{q}_h \in [L^2(\Omega)]^d :\ \boldsymbol{q}_h|_T \in [\mathbb{P}_k(T)]^d,\ \forall\ T \in \mathcal{T}_h \right\}.
\end{aligned}$$

Then the HDG method for (1.1) reads as follows. For all $(v_h, \widehat{v}_h, \boldsymbol{r}_h) \in V_h \times \widehat{V}_h^0 \times \boldsymbol{Q}_h$, find $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ such that

$$\begin{cases}
(a^{-1}\boldsymbol{p}_h, \boldsymbol{r}_h) + (u_h, \nabla_h \cdot \boldsymbol{r}_h) - \langle \widehat{u}_h, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \\
-(v_h, \nabla_h \cdot \boldsymbol{p}_h) + \langle \widehat{v}_h, \boldsymbol{p}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
+ \left\langle \tau \left(\Pi_k^\partial u_h - \widehat{u}_h, \Pi_k^\partial v_h - \widehat{v}_h \right) \right\rangle_{\partial \mathcal{T}_h} = -(f, v_h) + \langle g_N, \widehat{v}_h \rangle_{\Gamma_N},
\end{cases} \tag{2.4}$$

where

$$\tau|_E = W_E h_E^{-1}, \quad \forall E \in \mathcal{E}_h,$$

and W_E satisfies the following property:

$$W_E \sim \max_{E \subset \partial T} \{a_T\}.$$

REMARK 2.5 There are two simple choices for W_E:

(1)
$$W_E = a_T, \text{ when } E \subset \partial T \cap \partial \Omega,$$
$$W_E = \frac{a_{T_+} + a_{T_-}}{2}, \text{ when } E \text{ is shared by } T+ \text{ and } T- .$$

(2)
$$W_E = \max_{E \subset \partial T} \{a_T\}, \quad \forall E \subset \mathcal{E}_h.$$

To simplify notation, we define

$$\begin{aligned} B_h(\boldsymbol{p}, u, \widehat{u}; \boldsymbol{q}, v, \widehat{v}) &= (a^{-1}\boldsymbol{p}, \boldsymbol{q}) + (u, \nabla_h \cdot \boldsymbol{q}) - \langle \widehat{u}, \boldsymbol{q} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\ &\quad - (v, \nabla_h \cdot \boldsymbol{p}) + \langle \widehat{v}, \boldsymbol{p} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} + \langle \tau(\Pi_k^\partial u - \widehat{u}), \Pi_k^\partial v - \widehat{v} \rangle_{\partial \mathcal{T}_h}. \end{aligned} \tag{2.5}$$

Then (2.4) can be rewritten in a compact form as follows: find $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ such that

$$B_h(\boldsymbol{p}_h, u_h, \widehat{u}_h; \boldsymbol{r}_h, v_h, \widehat{v}_h) = -(f, v_h) + \langle g_N, \widehat{v}_h \rangle_{\Gamma_N}, \tag{2.6}$$

for all $(v_h, \widehat{v}_h, \boldsymbol{r}_h) \in V_h \times \widehat{V}_h^0 \times \boldsymbol{Q}_h$. Let $(u, \boldsymbol{p}) \in H^1(\Omega) \times \boldsymbol{H}(\mathrm{div}, \Omega)$ be the solution of (1.1). It follows from the definition of B_h and integration by parts that

$$B_h(\boldsymbol{p}, u, u; \boldsymbol{r}_h, v_h, \widehat{v}_h) = -(f, v_h) + \langle g_N, \widehat{v}_h \rangle_{\Gamma_N}, \tag{2.7}$$

for all $(\boldsymbol{r}_h, v_h, \widehat{v}_h) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h^0$.

By (2.6) and (2.7), we have the following orthogonality result.

THEOREM 2.6 (Orthogonality). Let $(u, \boldsymbol{p}) \in H^1(\Omega) \times \boldsymbol{H}(\mathrm{div}, \Omega)$ be the solution of (1.1) and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). Then for all $(v_h, \widehat{v}_h, \boldsymbol{r}_h) \in V_h \times \widehat{V}_h^0 \times \boldsymbol{Q}_h$, we have

$$B_h(\boldsymbol{p} - \boldsymbol{p}_h, u - u_h, u - \widehat{u}_h; \boldsymbol{r}_h, v_h, \widehat{v}_h) = 0. \tag{2.8}$$

2.3 *Projections*

To establish error estimates for the proposed HDG method, we need the following approximation and stability results for the L^2-projections Π_j^o and Π_j^∂ with nonnegative integer j.

LEMMA 2.7 Let m be an integer with $1 \leqslant m \leqslant j+1$. For all $T \in \mathscr{T}_h$, $E \in \mathscr{E}_h$, it holds that

$$\|\Pi_j^o v\|_{0,T} \leqslant \|v\|_{0,T} \qquad \forall\, v \in L^2(T), \tag{2.9}$$

$$\|\Pi_j^\partial v\|_{0,E} \leqslant \|v\|_{0,E} \qquad \forall\, v \in L^2(E), \tag{2.10}$$

$$\|v - \Pi_j^\partial v\|_{0,\partial T} \lesssim h_T^{m-1/2}|v|_{m,T} \qquad \forall\, v \in H^m(T), \tag{2.11}$$

$$|v - \Pi_j^o v|_{s,T} \lesssim h_T^{m-s}|v|_{m,T} \qquad \forall\, v \in H^m(T),\ 0 \leqslant s \leqslant m, \tag{2.12}$$

$$\|\nabla^s(v - \Pi_j^o v)\|_{0,\partial T} \lesssim h_T^{m-s-1/2}|v|_{m,T} \qquad \forall\, v \in H^m(T),\ 1 \leqslant s+1 \leqslant m, \tag{2.13}$$

where s is an integer.

Proof. The stability results (2.9–2.10) follow from the definitions of L^2-projections. The approximation result (2.13) follows directly from (2.2) and (2.12). Since $\|v - \Pi_j^\partial v\|_{0,\partial T} \leqslant \|v - \Pi_j^o v\|_{0,\partial T}$, the estimate (2.11) follows from (2.13) with $s = 0$. It only remains to prove (2.12). In fact, by combining (2.3), the inverse estimate (2.1) and the stability estimate (2.9), we have

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

□

LEMMA 2.8 For any nonnegative integer j, it holds that

$$\|v - \Pi_j^o v\|_{0,\partial T} \lesssim h_T^{1/2}|v - \Pi_j^o v|_{1,T}. \tag{2.14}$$

Proof. By the trace inequality (2.2) and the approximation property (2.12), we get

$$\begin{aligned}
\|v - \Pi_k^o v\|_{0,\partial T} &\lesssim h_T^{-1/2}\|v - \Pi_k^o v\|_{0,T} + h_T^{1/2}|v - \Pi_k^o v|_{1,T} \\
&= h_T^{-1/2}\|v - \Pi_k^o v - \Pi_k^o(v - \Pi_k^o v)\|_{0,T} + h_T^{1/2}|v - \Pi_k^o v|_{1,T} \\
&\lesssim h_T^{1/2}|v - \Pi_k^o v|_{1,T}.
\end{aligned}$$

□

Next, we recall the following classical results.

THEOREM 2.9 (Adams & Fournier, 2003, Page 220, Theorem 7.23 and Brenner & Scott, 2008, Page 373, Proposition 14.1.5) Given Banach spaces $A_1 \hookrightarrow A_0$ and $B_1 \hookrightarrow B_0$, let $\mathscr{K}\colon (A_0+A_1) \to (B_0+B_1)$

be a bounded linear operator from A_i into B_i ($i = 0, 1$). Then $\mathscr{K} : A_{\theta,p} \to B_{\theta,p}$ is a bounded linear operator for any $0 < \theta < 1$, $1 \leqslant p \leqslant \infty$. Moreover,

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

where $A_{\theta,p} := [A_0, A_1]_{\theta,p}$, $B_{\theta,p} := [B_0, B_1]_{\theta,p}$. See Brenner & Scott (2008, Page 372) for detailed definitions of $A_{\theta,p}$ and $B_{\theta,p}$.

THEOREM 2.10 (Brenner & Scott, 2008, Page 375, Theorem 14.2.7) If Ω has a Lipschitz boundary, then

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

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

With the above results, we are ready to derive the following fractional approximation properties of the L^2-projection Π_j^o.

LEMMA 2.11 Let j be a nonnegative integer, and let real numbers α, β satisfy $0 \leqslant \alpha < \beta \leqslant j+1$. Then for all $v \in H^\beta(T)$ and $T \in \mathscr{T}_h$,

$$\left\|\left(Id - \Pi_j^o\right)v\right\|_{\alpha,T} \lesssim h_T^{\beta-\alpha}\|v\|_{\beta,T},$$

where Id is the identity operator.

Proof. Let $r \geqslant \beta$ be an integer. When $r-1 < \alpha < r$, we take $A_0 = A_1 = H^r(\Omega)$, $B_0 = H^{r-1}(\Omega)$, $B_1 = H^r(\Omega)$ and $\theta = \alpha + 1 - r$ in Theorem 2.9. Then by combining (2.12) and Theorems 2.9 and 2.10, we have

$$\begin{aligned}\frac{\left\|\left(Id - \Pi_j^o\right)v\right\|_{\alpha,T}}{\|v\|_{r,T}} &\leqslant \left\|\left(Id - \Pi_j^o\right)\right\|_{H^r(T)\to H^\alpha(T)}\\ &\leqslant \left\|\left(Id - \Pi_j^o\right)\right\|_{H^r(T)\to H^{r-1}(T)}^{1-\theta}\left\|\left(Id - \Pi_j^o\right)\right\|_{H^r(T)\to H^r(T)}^{\theta}\\ &\lesssim h_T^{1-\theta}\\ &= h_T^{r-\alpha}.\end{aligned}$$

When $\alpha = r-1$ or $\alpha = r$, it follows from (2.12) that

$$\left\|\left(Id - \Pi_j^o\right)v\right\|_{\alpha,T} \lesssim h_T^{r-\alpha}\|v\|_{r,T}. \tag{2.15}$$

Therefore, (2.15) holds for $r-1 \leqslant \alpha \leqslant r$. In view of (2.12) and (2.15), we have

$$\begin{aligned}\left\|\left(Id - \Pi_j^o\right)v\right\|_{\alpha,T} &= \left\|\left(Id - \Pi_j^o\right)v - \Pi_j^o\left(Id - \Pi_j^o\right)v\right\|_{\alpha,T}\\ &\lesssim h_T^{1-\alpha}\left\|\left(Id - \Pi_j^o\right)v\right\|_{0,T}\\ &\lesssim h_T^{r-\alpha}\|v\|_{r,T}.\end{aligned}$$

In particular, when $r = \beta$ we have

$$\left\| \left(Id - \Pi_j^o \right) v \right\|_{\alpha,T} \lesssim h_T^{\beta-\alpha} \|v\|_{\beta,T}.$$

Next, we assume $\alpha < \beta < r$ and take $A_0 = H^\alpha(\Omega)$, $A_1 = H^r(\Omega)$, $B_0 = B_1 = H^r(\Omega)$, $\theta = \frac{\beta-\alpha}{r-\alpha}$. From (2.12), Theorems 2.9 and 2.10 and (2.15), we have

$$\begin{aligned}
\frac{\left\| \left(Id - \Pi_j^o \right) v \right\|_{\alpha,T}}{\|v\|_{\beta,T}} &\leqslant \left\| \left(Id - \Pi_j^o \right) \right\|_{H^\beta(T) \to H^\alpha(T)} \\
&\leqslant \left\| \left(Id - \Pi_j^o \right) \right\|_{H^\alpha(T) \to H^\alpha(T)}^{1-\theta} \left\| \left(Id - \Pi_j^o \right) \right\|_{H^r(T) \to H^\alpha(T)}^{\theta} \\
&\lesssim h_T^{(r-\alpha)\theta} \\
&= h_T^{\beta-\alpha},
\end{aligned}$$

which completes the proof. □

3. *A priori* error estimates

LEMMA 3.1 Let $(u, \boldsymbol{p}) \in H^1(\Omega) \times \boldsymbol{H}(\mathrm{div}, \Omega)$ be the solution of (1.1) and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). The following estimate holds:

$$a_T^{1/2} |u - u_h|_{1,T} \lesssim a_T^{-1/2} \left\| \boldsymbol{p} - \boldsymbol{p}_h \right\|_{0,T} + a_T^{1/2} h_T^{-1/2} \left\| \Pi_k^\partial u_h - \widehat{u}_h \right\|_{0,\partial T}, \tag{3.1}$$

for any $T \in \mathscr{T}_h$.

Proof. We apply integration by parts to the first equation of (2.4) and use the definition of Π_k^∂ to get

$$\left(a^{-1} \boldsymbol{p}_h, \boldsymbol{r}_h \right) - (\nabla_h u_h, \boldsymbol{r}_h) - \left\langle \widehat{u}_h - \Pi_k^\partial u_h, \boldsymbol{r}_h \cdot \boldsymbol{n} \right\rangle_{\partial \mathscr{T}_h} = 0.$$

For any $T \in \mathscr{T}_h$, we take $\boldsymbol{r}_h = a^{-1}\boldsymbol{p}_h - \nabla_h u_h \in [\mathbb{P}_k(T)]^d$ and $\boldsymbol{r}_h = 0$ on Ω / T in the equation above and use an inverse inequality to get

$$\begin{aligned}
\left\| a_T^{-1} \boldsymbol{p}_h - \nabla u_h \right\|_{0,T}^2 &= \left\langle \widehat{u}_h - \Pi_k^\partial u_h, \left(a^{-1} \boldsymbol{p}_h - \nabla u_h \right) \cdot \boldsymbol{n} \right\rangle_{\partial T} \\
&\lesssim \left\| a_T^{-1} \boldsymbol{p}_h - \nabla u_h \right\|_{0,T} h_T^{-1/2} \left\| \Pi_k^\partial u_h - \widehat{u}_h \right\|_{0,\partial T},
\end{aligned}$$

which leads to

$$\left\| a_T^{-1} \boldsymbol{p}_h - \nabla u_h \right\|_{0,T} \lesssim h_T^{-1/2} \left\| \Pi_k^\partial u_h - \widehat{u}_h \right\|_{0,\partial T}. \tag{3.2}$$

By multiplying $a_T^{1/2}$ to both sides of (3.2), we get

$$\left\| a_T^{-1/2} \boldsymbol{p}_h - a_T^{1/2} \nabla u_h \right\|_{0,T} \lesssim a_T^{1/2} h_T^{-1/2} \left\| \Pi_k^\partial u_h - \widehat{u}_h \right\|_{0,\partial T}. \tag{3.3}$$

Note that $\boldsymbol{p} = a\nabla u$. It then follows from (3.3) and the triangle inequality that

$$\begin{aligned} a_T^{1/2}\|\nabla u - \nabla u_h\|_{0,T} &= \left\| a_T^{-1/2}\boldsymbol{p} - a_T^{1/2}\nabla u_h \right\|_{0,T} \\ &\leqslant a_T^{-1/2}\left\|\boldsymbol{p} - \boldsymbol{p}_h\right\|_{0,T} + \left\| a_T^{-1/2}\boldsymbol{p}_h - a_T^{1/2}\nabla u_h \right\|_{0,T} \\ &\lesssim a_T^{-1/2}\left\|\boldsymbol{p} - \boldsymbol{p}_h\right\|_{0,T} + a_T^{1/2}h_T^{-1/2}\left\| \Pi_k^{\partial} u_h - \widehat{u}_h \right\|_{0,\partial T}, \end{aligned}$$

which proves the estimate (3.1). □

Lemma 3.2 Let $(u,\boldsymbol{p}) \in H^1(\Omega) \times \boldsymbol{H}(\mathrm{div},\Omega)$ be the solution of (1.1), and $(u_h,\widehat{u}_h,\boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). Then

$$\|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_0^2 + \left\| \tau^{1/2}\left(\Pi_k^{\partial} u_h - \widehat{u}_h\right)\right\|_{0,\partial\mathcal{T}_h}^2 = E(u,\boldsymbol{p};u,\widehat{u}_h,\boldsymbol{p}_h), \tag{3.4}$$

where

$$\begin{aligned} E(u,\boldsymbol{p};u,\widehat{u}_h,\boldsymbol{p}_h) = &\left(a^{-1}(\boldsymbol{p}-\boldsymbol{p}_h),\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right) - \left(\nabla_h(u-u_h),\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) \\ &+ \left(\nabla_h\left(u - \Pi_{k+1}^o u\right),\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) - \left\langle \Pi_k^{\partial}u_h - \widehat{u}_h, \left(\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right)\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\ &+ \left\langle \Pi_{k+1}^o(u-u_h) - \Pi_k^{\partial}(u-u_h), \left(\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right)\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\ &- \left\langle \tau\left(\Pi_k^{\partial}u_h - \widehat{u}_h\right), \Pi_k^{\partial}u - \Pi_{k+1}^o u\right\rangle_{\partial\mathcal{T}_h}, \end{aligned} \tag{3.5}$$

and $r = \min\{k,m\}$. Here, m is the integer part of s.

Proof. By the orthogonality of Π_k^{∂}, we have

$$\begin{aligned} \left\|\Pi_k^{\partial}(u-u_h) - (u-\widehat{u}_h)\right\|_{0,E}^2 &= \left\|\left(\Pi_k^{\partial}u - u\right) - \left(\Pi_k^{\partial}u_h - \widehat{u}_h\right)\right\|_{0,E}^2 \\ &= \left\|\Pi_k^{\partial}u - u\right\|_{0,E}^2 + \left\|\Pi_k^{\partial}u_h - \widehat{u}_h\right\|_{0,E}^2. \end{aligned} \tag{3.6}$$

It follows from (3.6), the definition of B_h and Lemma 2.1 that

$$\begin{aligned} &\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \left\|\tau^{1/2}\left(\Pi_k^{\partial}u_h - \widehat{u}_h\right)\right\|_{0,\partial\mathcal{T}_h}^2 + \left\|\tau^{1/2}\left(\Pi_k^{\partial}u - u\right)\right\|_{0,\partial\mathcal{T}_h}^2 \\ &= \|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \left\|\tau^{1/2}\left(\Pi_k^{\partial}(u-u_h) - (u-\widehat{u}_h)\right)\right\|_{0,\partial\mathcal{T}_h}^2 \\ &= B_h\left(\boldsymbol{p}-\boldsymbol{p}_h, u-u_h, u-\widehat{u}_h; \boldsymbol{p}-\boldsymbol{p}_h, u-u_h, u-\widehat{u}_h\right) \\ &= B_h\left(\boldsymbol{p}-\boldsymbol{p}_h, u-u_h, u-\widehat{u}_h; \boldsymbol{p}-\Pi_r^o\boldsymbol{p}, u-\Pi_{k+1}^o u, u-\Pi_k^{\partial}u\right), \end{aligned} \tag{3.7}$$

where we used the fact that $\widehat{u}_h - \Pi_k^\partial u = 0$ on Γ_D. In view of the definition of B_h in (2.5), we have

$$\begin{aligned}
&B_h\left(\boldsymbol{p}-\boldsymbol{p}_h, u-u_h, u-\widehat{u}_h; \boldsymbol{p}-\Pi_r^o\boldsymbol{p}, u-\Pi_{k+1}^o u, u-\Pi_k^\partial u\right)\\
&\quad= \left(a^{-1}(\boldsymbol{p}-\boldsymbol{p}_h), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) + \left(u-u_h, \nabla_h\cdot(\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\right) - \langle u-\widehat{u}_h, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}\\
&\quad- \left(u-\Pi_{k+1}^o u, \nabla_h\cdot(\boldsymbol{p}-\boldsymbol{p}_h)\right) + \left\langle u-\Pi_k^\partial u, (\boldsymbol{p}-\boldsymbol{p}_h)\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}\\
&\quad+ \left\langle \tau\left(\Pi_k^\partial(u-u_h)-(u-\widehat{u}_h)\right), \Pi_k^\partial\left(u-\Pi_{k+1}^o u\right) - \left(u-\Pi_k^\partial u\right)\right\rangle_{\partial\mathscr{T}_h}\\
&\quad=: \sum_1^6 R_i.
\end{aligned} \tag{3.8}$$

Integrating by parts gives

$$\begin{aligned}
R_2+R_3 &= -\left(\nabla_h(u-u_h), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) - \langle u_h-\widehat{u}_h, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}\\
&= -\left(\nabla_h(u-u_h), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) - \left\langle \Pi_k^\partial u_h-\widehat{u}_h, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}\\
&\quad- \left\langle u_h-\Pi_k^\partial u_h, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}.
\end{aligned} \tag{3.9}$$

Using integration by parts and the orthogonality of Π_k^∂, we arrive at

$$R_4 = \left(\nabla_h\left(u-\Pi_{k+1}^o u\right), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) - \langle u-\Pi_{k+1}^o u, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}, \tag{3.10}$$

$$R_5 = \left\langle u-\Pi_k^\partial u, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}. \tag{3.11}$$

Then from (3.9, 3.10–3.11), we have

$$\begin{aligned}
R_2+R_3+R_4+R_5 &= -\left(\nabla_h(u-u_h), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right) + \left(\nabla_h\left(u-\Pi_{k+1}^o u\right), \boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right)\\
&\quad- \left\langle \Pi_k^\partial u_h-\widehat{u}_h, (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}\\
&\quad+ \left\langle \Pi_{k+1}^o(u-u_h)-\Pi_k^\partial(u-u_h), (\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}.
\end{aligned} \tag{3.12}$$

It follows from the orthogonality and the definition of Π_k^∂ that

$$\begin{aligned}
R_6 &= \left\langle \tau\left(\Pi_k^\partial u-u\right), \Pi_k^\partial\left(u-\Pi_{k+1}^o u\right) - \left(u-\Pi_k^\partial u\right)\right\rangle_{\partial\mathscr{T}_h}\\
&\quad- \left\langle \tau\left(\Pi_k^\partial u_h-\widehat{u}_h\right), \Pi_k^\partial\left(u-\Pi_{k+1}^o u\right) - \left(u-\Pi_k^\partial u\right)\right\rangle_{\partial\mathscr{T}_h}\\
&= \left\|\tau^{1/2}\left(\Pi_k^\partial u-u\right)\right\|_{0,\partial\mathscr{T}_h}^2 - \left\langle \tau\left(\Pi_k^\partial u_h-\widehat{u}_h\right), \Pi_k^\partial u-\Pi_{k+1}^o u\right\rangle_{\partial\mathscr{T}_h}.
\end{aligned} \tag{3.13}$$

The desired result (3.4) is then proved by combining (3.7), (3.8), (3.12) and (3.13). □

THEOREM 3.3 Let $(u,\boldsymbol{p}) \in H^{1+s}(\Omega) \times \boldsymbol{H}(\mathrm{div},\Omega) \cap [H^s(\Omega)]^d$ (with $s > 0$) be the solution of (1.1), and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). It holds that

$$\begin{aligned}
&\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \|a^{1/2}(\nabla u - \nabla_h u_h)\|_0^2 + \|\tau^{1/2}(\Pi_k^\partial u_h - \widehat{u}_h)\|_{0,\partial\mathscr{T}_h}^2 \\
&\quad \lesssim \sum_{T\in\mathscr{T}_h} a_T |u - \Pi_{k+1}^o u|_{1,T}^2 + \sum_{T\in\mathscr{T}_h} a_T^{-1}\|\boldsymbol{p} - \Pi_r^o \boldsymbol{p}\|_{0,T}^2 + \sum_{T\in\mathscr{T}_h} a_T^{-1} h_T^2 \|\nabla\cdot\boldsymbol{p} - \nabla\cdot\Pi_r^o\boldsymbol{p}\|_{0,T}^2, \quad (3.14)
\end{aligned}$$

where $r = \min\{k, m\}$, and m is the integer part of s.

Proof. To simplify notation, we define

$$\begin{aligned}
E_1 &= \left(a^{-1}(\boldsymbol{p}-\boldsymbol{p}_h), \boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right),\\
E_2 &= -\left(\nabla_h(u-u_h), \boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right),\\
E_3 &= \left(\nabla_h\left(u - \Pi_{k+1}^o u\right), \boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right),\\
E_4 &= -\left\langle \Pi_k^\partial u_h - \widehat{u}_h, \left(\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right)\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h},\\
E_5 &= \left\langle \Pi_{k+1}^o(u-u_h) - \Pi_k^\partial(u-u_h), \left(\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\right)\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h},\\
E_6 &= -\left\langle \tau\left(\Pi_k^\partial u_h - \widehat{u}_h\right), \Pi_k^\partial u - \Pi_{k+1}^o\right\rangle_{\partial\mathscr{T}_h}.
\end{aligned}$$

It follows from the Cauchy–Schwarz inequality that

$$\begin{aligned}
E_1 &\leqslant \sum_{T\in\mathscr{T}_h} a_T^{-1}\|\boldsymbol{p}-\boldsymbol{p}_h\|_{0,T}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}\\
&\leqslant \|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0 \left(\sum_{T\in\mathscr{T}_h} a_T^{-1}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}^2\right)^{1/2}, \quad (3.15)\\
E_2 &\leqslant \sum_{T\in\mathscr{T}_h} \|\nabla(u-u_h)\|_{0,T}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}\\
&\leqslant \|a^{1/2}\nabla_h(u-u_h)\|_0 \left(\sum_{T\in\mathscr{T}_h} a_T^{-1}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}^2\right)^{1/2}, \quad (3.16)\\
E_3 &\leqslant \sum_{T\in\mathscr{T}_h} \|\nabla(u - \Pi_{k+1}^o u)\|_{0,T}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}\\
&\leqslant \left(\sum_{T\in\mathscr{T}_h} a_T h_T^{2s}\|u\|_{1+s,T}^2\right)^{1/2} \left(\sum_{T\in\mathscr{T}_h} a_T^{-1}\|\boldsymbol{p} - \Pi_r^o\boldsymbol{p}\|_{0,T}^2\right)^{1/2}. \quad (3.17)
\end{aligned}$$

Next we use the Cauchy–Schwarz inequality, the inverse inequality (2.1) and Lemma 2.3 with $\nabla v = \boldsymbol{p}$ to derive

$$
\begin{aligned}
E_4 &\lesssim \sum_{T\in\mathscr{T}_h}\sum_{E\subset\partial T} h_E^{-1/2}\left\|\Pi_k^{\partial}u_h-\widehat{u}_h\right\|_{0,E}\left(\|\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\|_{0,T}+h_T\|\nabla\cdot\left(\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\right)\|_{0,T}\right)\\
&\leqslant \left\|\tau^{1/2}\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right)\right\|_{0,\partial\mathscr{T}_h}\left(\sum_{T\in\mathscr{T}_h}a_T^{-1}\|\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\|_{0,T}^2+\sum_{T\in\mathscr{T}_h}a_T^{-1}h_T^2\|\nabla\cdot\boldsymbol{p}-\nabla\cdot\Pi_r^o\boldsymbol{p}\|_{0,T}^2\right)^{1/2}.
\end{aligned}
\tag{3.18}
$$

From the Cauchy–Schwarz inequality, Lemma 2.3 and the approximation properties of Π_{k+1}^o and Π_k^{∂}, we obtain

$$
\begin{aligned}
E_5 &\lesssim \sum_{T\in\mathscr{T}_h}\sum_{E\subset\partial T} h_T^{-1/2}\|\Pi_{k+1}^o(u-u_h)-\Pi_k^{\partial}(u-u_h)\|_{0,E}\left(\|\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\|_{0,T}+h_T\|\nabla\cdot(\boldsymbol{p}-\Pi_r^o\boldsymbol{p})\|_{0,T}\right)\\
&\leqslant \|a^{1/2}(\nabla u-\nabla_h u_h)\|_0\left(\sum_{T\in\mathscr{T}_h}a_T^{-1}\|\boldsymbol{p}-\Pi_r^o\boldsymbol{p}\|_{0,T}^2+\sum_{T\in\mathscr{T}_h}a_T^{-1}h_T^2\|\nabla\cdot\boldsymbol{p}-\nabla\cdot\Pi_r^o\boldsymbol{p}\|_{0,T}^2\right)^{1/2}.
\end{aligned}
\tag{3.19}
$$

Similarly, we have

$$
\begin{aligned}
E_6 &= -\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right), u-\Pi_{k+1}^o u-\Pi_{k+1}^o\left(u-\Pi_{k+1}^o u\right)\right\rangle_{\partial\mathscr{T}_h}\\
&\lesssim \sum_{E\in\mathscr{E}_h} W_E h_E^{-1}\|\Pi_k^{\partial}u_h-\widehat{u}_h\|_{0,E}\left\|u-\Pi_{k+1}^o u-\Pi_{k+1}^o\left(u-\Pi_{k+1}^o u\right)\right\|_{0,E}\\
&\lesssim \sum_{E\in\mathscr{E}_h} W_E h_E^{-1}\|\Pi_k^{\partial}u_h-\widehat{u}_h\|_{0,E}h_E^{1/2}|u-\Pi_0^o u|_{1,T_{\max}}\\
&\lesssim \left\|\tau^{1/2}\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right)\right\|_{0,\partial T_h}\left(\sum_{T\in\mathscr{T}_h}a_T|u-\Pi_{k+1}^o u|_{1,T}^2\right)^{1/2},
\end{aligned}
\tag{3.20}
$$

where $E\subset\partial T_{\max}$, and $T_{\max}\in\mathscr{T}_h$ is the element such that $a_{T_{\max}}=\max_{E\subset\partial T}a_T$. The desired estimate (3.14) follows from Lemma 3.2 and (3.15–3.20). □

We can then derive the *a priori* error estimates in the next theorem.

THEOREM 3.4 Let $(u,\boldsymbol{p}) \in H^{1+s}(\Omega) \times \boldsymbol{H}(\mathrm{div},\Omega) \cap [H^s(\Omega)]^d$ (with $0 < s < 1$) be the solution of (1.1), and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). It holds that

$$\begin{aligned}\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \|a^{1/2}(\nabla u - \nabla_h u_h)\|_0^2 + \left\|\tau^{1/2}\left(\Pi_k^\partial u_h - \widehat{u}_h\right)\right\|_{0,\partial\mathcal{T}_h}^2 \\ \lesssim \sum_{T\in\mathcal{T}_h} a_T h_T^{2s}\|u\|_{1+s,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1} h_T^{2s}\|\boldsymbol{p}\|_{s,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1} h_T^2 \|f\|_{0,T}^2. \qquad (3.21)\end{aligned}$$

Proof. Since $s \in (0,1)$, we have $r = 0$ in Theorem 3.3. Therefore,

$$\begin{aligned}\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \|a^{1/2}(\nabla u - \nabla_h u_h)\|_0^2 + \left\|\tau^{1/2}\left(\Pi_k^\partial u_h - \widehat{u}_h\right)\right\|_{0,\partial\mathcal{T}_h}^2 \\ \lesssim \sum_{T\in\mathcal{T}_h} a_T |u - \Pi_{k+1}^o u|_{1,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1}\|\boldsymbol{p} - \Pi_0^o \boldsymbol{p}\|_{0,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1} h_T^2 \|\nabla\cdot\boldsymbol{p}\|_{0,T}^2. \qquad (3.22)\end{aligned}$$

Since $\nabla\cdot\boldsymbol{p} = f$, we can obtain the error estimate (3.21) directly from Lemma 2.11:

$$\begin{aligned}\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2 + \|a^{1/2}(\nabla u - \nabla_h u_h)\|_0^2 + \left\|\tau^{1/2}\left(\Pi_k^\partial u_h - \widehat{u}_h\right)\right\|_{0,\partial\mathcal{T}_h}^2 \\ \lesssim \sum_{T\in\mathcal{T}_h} a_T h_T^{2s}\|u\|_{1+s,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1} h_T^{2s}\|\boldsymbol{p}\|_{s,T}^2 + \sum_{T\in\mathcal{T}_h} a_T^{-1} h_T^2 \|f\|_{0,T}^2.\end{aligned}$$

□

4. *A posteriori* error estimates

We first introduce the following oscillation terms:

$$\mathrm{osc}^2(f,\mathcal{T}_h) = \sum_{T\in\mathcal{T}_h}\mathrm{osc}^2(f,T) = \sum_{T\in\mathcal{T}_h} a_T^{-1}h_T^2\|f - \Pi_{k+1}^o f\|_{0,T}^2,$$

$$\mathrm{osc}^2\left(g_N,\mathcal{E}_h^N\right) = \sum_{E\in\mathcal{E}_h^N}\mathrm{osc}^2(g_N,E) = \sum_{E\in\mathcal{E}_h^N} a_E^{-1}h_E\|g_N - \Pi_k^\partial g_N\|_{0,E}^2.$$

Then we define the *a posteriori* error estimators as follows:

$$\eta_{r_1,T} = a_T^{1/2}\|\nabla u_h - a^{-1}\boldsymbol{p}_h\|_{0,T}, \qquad (4.1)$$

$$\eta_{r_2,T} = a_T^{-1/2}h_T\|\Pi_{k+1}^o f - \nabla\cdot\boldsymbol{p}_h\|_{0,T}, \qquad (4.2)$$

$$\eta_{u_s,E} = W_E^{1/2}h_E^{-1/2}\|\Pi_k^\partial u_h - \widehat{u}_h\|_{0,E}, \qquad (4.3)$$

$$\eta_{u_j,E} = \begin{cases} W_E^{1/2} h_E^{-1/2} \left\| \left[\!\left[\left(Id - \Pi_k^{\partial} \right) u_h \right]\!\right] \right\|_{0,E}, \ E \in \mathscr{E}_h^I, \\ W_E^{1/2} h_E^{-1/2} \left\| \left(Id - \Pi_k^{\partial} \right) (u_h - g_D) \right\|_{0,E}, \ E \in \mathscr{E}_h^D, \\ 0, \ E \in \mathscr{E}_h^N, \end{cases} \tag{4.4}$$

and

$$\eta_{r_i}^2 = \sum_{T \in \mathscr{T}_h} \eta_{r_i,T}^2, \qquad i = 1, 2, \tag{4.5}$$

$$\eta_{u_s}^2 = \sum_{E \in \mathscr{E}_h} \eta_{u_s,E}^2, \tag{4.6}$$

$$\eta_{u_j}^2 = \sum_{E \in \mathscr{E}_h^I} \eta_{u_j,E}^2 + \sum_{E \in \mathscr{E}_h^D} \eta_{u_j,E}^2, \tag{4.7}$$

$$\eta^2 = \eta_{r_1}^2 + \eta_{r_2}^2 + \eta_{u_s}^2 + \eta_{u_j}^2 + \mathrm{osc}^2(f, \mathscr{T}_h) + \mathrm{osc}^2\left(g_N, \mathscr{E}_h^N\right). \tag{4.8}$$

Note that there are no explicit oscillation terms for g_D in constructing the global *a posteriori* error estimator η. Actually, those oscillation terms are involved implicitly by introducing (4.4).

4.1 *Reliability*

THEOREM 4.1 (Upper bound). Let $(u, \boldsymbol{p}) \in H^{1+s}(\Omega) \times (\boldsymbol{H}(\mathrm{div}, \Omega) \cap [H^s(\Omega)]^d)$ with $s > 0$, be the solution of (1.1) and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). Then

$$\|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_0 + \|a^{1/2}(\nabla u - \nabla_h u_h)\|_0 \lesssim \eta. \tag{4.9}$$

Proof. Let $(\boldsymbol{\gamma}, w, \widehat{w}) = (\boldsymbol{p} - \boldsymbol{p}_h - \boldsymbol{r}_h, u - u_h - v_h, u - \widehat{u}_h - \widehat{v}_h)$, where

$$\boldsymbol{r}_h := \Pi_0^o(\boldsymbol{p} - \boldsymbol{p}_h) \in \boldsymbol{Q}_h, \tag{4.10}$$

$$v_h := \Pi_{k+1}^o(u - u_h) \in V_h, \tag{4.11}$$

$$\widehat{v}_h := \Pi_k^{\partial}(u - \widehat{u}_h) \in \widehat{V}_h^0. \tag{4.12}$$

Note that

$$\|\Pi_k^{\partial}(u - u_h) - (u - \widehat{u}_h)\|_{0,E}^2 = \left\| \left(\Pi_k^{\partial} u - u\right) - \left(\Pi_k^{\partial} u_h - \widehat{u}_h\right) \right\|_{0,E}^2 = \|\Pi_k^{\partial} u - u\|_{0,E} + \|\Pi_k^{\partial} u_h - \widehat{u}_h\|_{0,E}^2.$$

The above equation together with the definition of B_h implies

$$\|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_0^2 + \eta_{u_s}^2 = B_h(\boldsymbol{p} - \boldsymbol{p}_h, u - u_h, u - \widehat{u}_h; \boldsymbol{p} - \boldsymbol{p}_h, u - u_h, u - \widehat{u}_h) - \left\| \tau^{1/2} \left(Id - \Pi_k^{\partial}\right) u \right\|_{0,\partial \mathscr{T}_h}^2.$$

It then follows from Lemma 2.1 that

$$
\begin{aligned}
\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\eta_{u_s}^2 &= B_h(\boldsymbol{p}-\boldsymbol{p}_h,u-u_h,u-\widehat{u}_h;\boldsymbol{\gamma},w,\widehat{w})-\left\|\tau^{1/2}\left(Id-\Pi_k^{\partial}\right)u\right\|_{0,\mathscr{T}_h}^2\\
&=(a^{-1}(\boldsymbol{p}-\boldsymbol{p}_h),\boldsymbol{\gamma})+(u-u_h,\nabla_h\cdot\boldsymbol{\gamma})-\langle u-\widehat{u}_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}\\
&\quad-(w,\nabla\cdot\boldsymbol{p}-\nabla_h\cdot\boldsymbol{p}_h)+\langle\widehat{w},\boldsymbol{p}-\boldsymbol{p}_h\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}\\
&\quad-\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right),\Pi_k^{\partial}w-\widehat{w}\right\rangle_{\partial\mathscr{T}_h},
\end{aligned}
$$

where we have used the relation

$$
\begin{aligned}
&\langle\tau(\Pi_k^{\partial}(u-u_h)-(u-\widehat{u}_h)),\Pi_k^{\partial}w-\widehat{w}\rangle_{\partial\mathscr{T}_h}\\
&\quad=\left\langle\tau\left(\Pi_k^{\partial}u-u\right),\Pi_k^{\partial}w-\widehat{w}\right\rangle_{\partial\mathscr{T}_h}-\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right),\Pi_k^{\partial}w-\widehat{w}\right\rangle_{\partial\mathscr{T}_h}\\
&\quad=\left\|\tau^{1/2}\left(Id-\Pi_k^{\partial}\right)u\right\|_{0,\mathscr{T}_h}^2-\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right),\Pi_k^{\partial}w-\widehat{w}\right\rangle_{\partial\mathscr{T}_h}.
\end{aligned}
$$

By using integration by parts, the relations $a^{-1}\boldsymbol{p}=\nabla u$ and $\nabla\cdot\boldsymbol{p}=f$, and the facts $(w,\ \Pi_{k+1}^{o}f-\nabla_h\cdot\boldsymbol{p}_h)=0$ and $\langle\Pi_k^{\partial}u-u,\ \boldsymbol{\Pi}_0^{o}(\boldsymbol{p}-\boldsymbol{p}_h)\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}=0$, we get

$$
\begin{aligned}
\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\eta_{u_s}^2 &= \langle\widehat{u}_h-u_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}+\langle\widehat{w},(\boldsymbol{p}-\boldsymbol{p}_h)\cdot\boldsymbol{n}\rangle_{\partial\mathscr{T}_h}\\
&\quad-\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right),w\right\rangle_{\partial\mathscr{T}_h}+\left(\nabla_hu_h-a^{-1}\boldsymbol{p}_h,\boldsymbol{\gamma}\right)+\left(w,f-\Pi_{k+1}^{o}f\right)\\
&=\left\langle\widehat{u}_h-u_h+u-\Pi_k^{\partial}u,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}-\left\langle\tau\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right),w\right\rangle_{\partial\mathscr{T}_h}\\
&\quad+\left(\nabla_hu_h-a^{-1}\boldsymbol{p}_h,\boldsymbol{\gamma}\right)+\left(w,f-\Pi_{k+1}^{o}f\right)\\
&=:R_1+R_2+R_3+R_4.
\end{aligned}
$$

Next, we estimate each R_i term by term. By the orthogonality of Π_k^{∂}, we have

$$
\begin{aligned}
R_1&=\left\langle\widehat{u}_h-\Pi_k^{\partial}u_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}+\left\langle\Pi_k^{\partial}u_h-u_h+u-\Pi_k^{\partial}u,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}\\
&=\left\langle\widehat{u}_h-\Pi_k^{\partial}u_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}+\left\langle\Pi_k^{\partial}u_h-u_h+u-\Pi_k^{\partial}u,\boldsymbol{p}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}\\
&=\left\langle\widehat{u}_h-\Pi_k^{\partial}u_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}+\left\langle[\![(\Pi_k^{\partial}-Id)u_h]\!],\boldsymbol{p}\cdot\boldsymbol{n}\right\rangle_{\mathscr{E}_h^I}+\left\langle\left(\Pi_k^{\partial}-Id\right)(u_h-g_D),\boldsymbol{p}\cdot\boldsymbol{n}\right\rangle_{\mathscr{E}_h^D}\\
&\quad+\left\langle\left(\Pi_k^{\partial}-Id\right)(u_h-u),g_N-\Pi_k^{\partial}g_N\right\rangle_{\mathscr{E}_h^N}\\
&=\left\langle\widehat{u}_h-\Pi_k^{\partial}u_h,\boldsymbol{\gamma}\cdot\boldsymbol{n}\right\rangle_{\partial\mathscr{T}_h}+\left\langle[\![(\Pi_k^{\partial}-Id)u_h]\!],(\boldsymbol{p}-\{\!\{\boldsymbol{p}_h\}\!\})\cdot\boldsymbol{n}\right\rangle_{\mathscr{E}_h^I}\\
&\quad+\left\langle\left(\Pi_k^{\partial}-Id\right)(u_h-g_D),(\boldsymbol{p}-\boldsymbol{p}_h)\cdot\boldsymbol{n}\right\rangle_{\mathscr{E}_h^D}+\left\langle\left(\Pi_k^{\partial}-Id\right)(u_h-u),g_N-\Pi_k^{\partial}g_N\right\rangle_{\mathscr{E}_h^N}.
\end{aligned}
$$

Then by the Cauchy–Schwarz inequality and Lemma 2.3, we get

$$
\begin{aligned}
R_1 &\lesssim (\eta_{u_s}+\eta_{u_j})\left(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_2}+\operatorname{osc}(f,\mathscr{T}_h)\right)+\operatorname{osc}(g_N,\mathscr{E}_h^N)\|a^{1/2}(\nabla u-\nabla_h u_h)\|_0\\
&\lesssim (\eta_{u_s}+\eta_{u_j})\left(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_2}+\operatorname{osc}(f,\mathscr{T}_h)\right)+\operatorname{osc}(g_N,\mathscr{E}_h^N)\left(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_1}\right).
\end{aligned}
$$

Using Lemma 2.3, we have

$$
\begin{aligned}
R_2 &\leqslant \sum_{T\in\mathscr{T}_h}\sum_{E\subset\partial T}\left\|\tau^{1/2}\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right)\right\|_{0,E}W_E^{1/2}h_E^{-1/2}\|w\|_{0,E}\\
&\lesssim \sum_{T\in\mathscr{T}_h}\sum_{E\subset\partial T}\left\|\tau^{1/2}\left(\Pi_k^{\partial}u_h-\widehat{u}_h\right)\right\|_{E}W_E^{1/2}\|\nabla u-\nabla u_h\|_{0,T}\\
&\lesssim \eta_{u_s}\|a^{1/2}(\nabla u-\nabla u_h)\|_0\\
&\lesssim \eta_{u_s}\left(\eta_{r_1}+\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0\right),\\
R_3 &\leqslant \sum_{T\in\mathscr{T}_h}\|\nabla_h u_h-a^{-1}\boldsymbol{p}_h\|_{0,T}\|\boldsymbol{p}-\boldsymbol{p}_h\|_{0,T}\lesssim \eta_{r_1}\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0,\\
R_4 &\lesssim \|a^{1/2}(\nabla u-\nabla u_h)\|_0\cdot\operatorname{osc}(f,\mathscr{T}_h)\lesssim \operatorname{osc}(f,\mathscr{T}_h)(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_1}).
\end{aligned}
$$

By combining the above estimates for R_i, $i=1,2,3,4$, we arrive at

$$
\begin{aligned}
&\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\eta_{u_s}^2\\
&\quad\lesssim (\eta_{u_s}+\eta_{u_j})\left(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_2}+\operatorname{osc}(f,\mathscr{T}_h)\right)+\operatorname{osc}(g_N,\mathscr{E}_h^N)\left(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_1}\right)\\
&\qquad+\eta_{u_s}\left(\eta_{r_1}+\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0\right)+\eta_{r_1}\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\operatorname{osc}(f,\mathscr{T}_h)(\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0+\eta_{r_1}).
\end{aligned}
\tag{4.13}
$$

Finally, the estimate (4.9) can be obtained in a straightforward way using Lemma 3.1 and (4.13). □

4.2 *Efficiency*

In the rest of this section, we show that the proposed *a posteriori* error estimators are also efficient, i.e. lower bounds hold.

For any $E\in\mathscr{E}_h$, we denote

$$\omega_E=\{\text{the union of } T: E\subset\partial T,\ T\in\mathscr{T}_h\}.$$

Let $b_M\in H_0^1(M)$ be the usual bubble function defined on $M\in\mathscr{M}_h$. The following result is a standard tool for the *a posteriori* error estimates.

LEMMA 4.2 (Verfürth, 1994). For every $M \in \mathscr{M}_h$, it holds that

$$\|v_h\|_{0,M} \lesssim \left\|b_M^{1/2} v_h\right\|_{0,M} \lesssim \|v_h\|_{0,M}, \quad \forall\, v_h \in \mathbb{P}_k(M),\ k \geqslant 0. \tag{4.14}$$

THEOREM 4.3 (Lower bounds). Let $(u, \boldsymbol{p}) \in H^{1+s}(\Omega) \times (\boldsymbol{H}(\mathrm{div}, \Omega) \cap [H^s(\Omega)]^d)$ (with $s > 0$) be the solution of (1.1) and $(u_h, \widehat{u}_h, \boldsymbol{p}_h) \in V_h \times \widehat{V}_h^{g_D} \times \boldsymbol{Q}_h$ be the solution of (2.6). Then for any $T \in \mathscr{T}_h$ and $E \in \mathscr{E}_h$, it holds that

$$\eta_{r_1,T} \leqslant \|a^{1/2}(\nabla u - \nabla u_h)\|_{0,T} + \|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_{0,T}, \tag{4.15}$$

$$\eta_{r_2,T} \lesssim \|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_{0,T} + \mathrm{osc}(f, T), \tag{4.16}$$

$$\eta_{u_j,E} \lesssim \|a^{1/2}(\nabla u - \nabla u_h)\|_{0,\omega_E}, \tag{4.17}$$

$$\eta_{u_s} \lesssim \|a^{1/2}(\nabla u - \nabla u_h)\|_0 + \|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_0 + \mathrm{osc}(f, \mathscr{T}_h) + \mathrm{osc}\left(g_N, \mathscr{E}_h^N\right). \tag{4.18}$$

Proof of (4.15). By using the triangle inequality and the fact $a^{-1}\boldsymbol{p} = \nabla u$, we have

$$\begin{aligned}
\eta_{r_1,T} &\leqslant a_T^{1/2} \|\nabla u - \nabla u_h\|_{0,T} + a_T^{-1/2} \|\boldsymbol{p} - \boldsymbol{p}_h\|_{0,T} \\
&= \|a^{1/2}(\nabla u - \nabla u_h)\|_{0,T} + \|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_{0,T}.
\end{aligned}$$

□

Proof of (4.16). For any $M \in \mathscr{M}_h$, we let $b_M \in H_0^1(M)$ be the bubble function. It then follows from Lemma 4.2, the relation $\nabla \cdot \boldsymbol{p} = f$, the triangle inequality and integration by parts that

$$\begin{aligned}
\eta_{r_2,T}^2 &\lesssim \sum_{M \subset T} (\eta_{r_2,T}, b_M \eta_{r_2,T})_M \\
&= a_T^{-1/2} h_T \sum_{M \subset T} (\nabla \cdot \boldsymbol{p} - \nabla \cdot \boldsymbol{p}_h, b_M \eta_{r_2,T})_M + a_T^{-1/2} h_T \sum_{M \subset T} \left(\Pi_{k+1}^o f - f, b_M \eta_{r_2,T}\right)_M \\
&= -a_T^{-1/2} h_T \sum_{M \subset T} (\boldsymbol{p} - \boldsymbol{p}_h, \nabla(b_M \eta_{r_2,T}))_M + a_T^{-1/2} h_T \sum_{M \subset T} \left(\Pi_{k+1}^o f - f, b_M \eta_{r_2,T}\right)_M.
\end{aligned}$$

Then by the Cauchy–Schwarz inequality and the inverse inequality (2.1), we have

$$\eta_{r_2,T}^2 \lesssim \left(\|a^{-1/2}(\boldsymbol{p} - \boldsymbol{p}_h)\|_{0,T} + \mathrm{osc}(f, T)\right) \eta_{r_2,T},$$

which leads to (4.16). □

Proof of (4.17). It directly follows from the definition of $\eta_{u_j,E}$ and $W_E \sim \max_{E\subset\partial T}\{a_T\}$ that

$$\begin{aligned}\eta^2_{u_j,E} &= W_E h_E^{-1}\left\|\left(Id-\Pi_k^{\partial}\right)(u_h-u)\right\|^2_{0,E}\\ &\leqslant \sum_{E\subset\partial T} a_T h_E^{-1}\left\|\left(Id-\Pi_k^{\partial}\right)(u_h-u)\right\|^2_{0,\partial T}\\ &\lesssim \sum_{E\subset\partial T} a_T\|(\nabla u-\nabla u_h)\|^2_{0,T}\\ &= \|a(\nabla u-\nabla u_h)\|^2_{0,\omega_E}.\end{aligned}$$

□

Proof of (4.18). By combining (4.13), (4.15), (4.16) and (4.17), we have

$$\begin{aligned}\eta^2_{u_s} &\lesssim \eta^2_{r_1}+\eta^2_{r_2}+\eta^2_{u_j}+\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\mathrm{osc}^2(f,\mathcal{T}_h)+\mathrm{osc}^2\left(g_N,\mathcal{E}_h^N\right)\\ &\lesssim \|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\|a^{1/2}(\nabla u-\nabla_h u_h)\|_0^2+\mathrm{osc}^2(f,\mathcal{T}_h)+\mathrm{osc}^2\left(g_N,\mathcal{E}_h^N\right).\end{aligned}$$

□

5. Numerical experiments

In this section, we present the results of numerical experiments in two dimensions to demonstrate the efficiency and reliability of the *a posteriori* estimators. All tests are programmed in C++ using the Eigen library (Eigen 3.2.5) and Hypre library (Falgout & Yang, 2002). The numerical results in this section are obtained by the following adaptive mesh refinement algorithm.

Let

$$\eta_T^2=\sum_{E\subset\partial T}\left(\eta^2_{u_s,E}+\eta^2_{u_j,E}\right)+\eta^2_{r_1,T}+\eta^2_{r_2,T}+\mathrm{osc}^2(f,T)+\sum_{E\subset\mathcal{E}_h^N\cap\partial T}\mathrm{osc}^2(g_N,E),$$

and $\quad e=\|a^{-1/2}(\boldsymbol{p}-\boldsymbol{p}_h)\|_0^2+\|a^{1/2}(\nabla u-\nabla_h u_h)\|_0^2.$

Adaptive Algorithm. Starting with an initial mesh $\mathcal{T}_l$ $(l=0)$, choose a parameter $\beta\in[0,1]$ and take the following iterative steps:

(i) Solve the discrete problem on $\mathcal{T}_l$ with N degrees of freedom.

(ii) Compute η_T for all $T\in\mathcal{T}_l$ and $\eta=\left(\sum_{T\in\mathcal{T}_l}\eta_T^2\right)^{1/2}$.

(iii) Mark a set of elements $\mathcal{R}_l\subset\mathcal{T}_l$ with minimum number of elements such that $\sum_{T\in\mathcal{R}_l}\eta_T^2\geqslant\beta\eta^2$.

(iv) Refine all the elements in $\mathcal{R}_l$ to get $\mathcal{T}_{l+1}$.

(v) Further refine the elements to ensure there is at most one hanging node per edge. Update $l=l+1$ and go to (i).

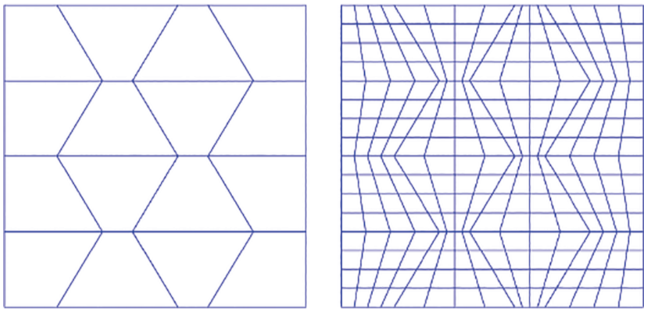

FIG. 2. Meshes for the smooth problem: levels 0 (left) and 2 (right) with $\beta = 1.0$.

TABLE 1 *Results for the smooth problem*

	$k=1$			$k=2$			$k=3$		
Mesh	e	η	η/e	e	η	η/e	e	η	η/e
0	6.0614E+00	8.1840E+00	1.35	3.5025E+00	3.9717E+00	1.13	1.4339E+00	1.5481E+00	1.08
1	2.5069E+00	3.2953E+00	1.31	8.3972E–01	9.6953E–01	1.15	1.8555E–01	2.0584E–01	1.11
2	1.1468E+00	1.5591E+00	1.36	1.9935E–01	2.4057E–01	1.21	2.2714E–02	2.5767E–02	1.13
3	5.5500E–01	7.7380E–01	1.39	4.8625E–02	6.0638E–02	1.25	2.7822E–03	3.2135E–03	1.16
4	2.7420E–01	3.8748E–01	1.41	1.2079E–02	1.5302E–02	1.27	3.4486E–04	4.0209E–04	1.17
5	1.3656E–01	1.9406E–01	1.42	3.0192E–03	3.8477E–03	1.27	4.3030E–05	5.0342E–05	1.17
6	6.8202E–02	9.7123E–02	1.42	7.5540E–04	9.6488E–04	1.28	5.3794E–06	6.2997E–06	1.17

5.1 *Smooth problem*

Consider the problem (1.1) on the unit square with $a = 1$, $\Gamma_N = \emptyset$, $g_D = 0$, and f is chosen according to the following exact solution:

$$u = \sin(2\pi x)\sin(2\pi y), \quad \text{and} \quad \boldsymbol{p} = \nabla u.$$

We take $\beta = 1.0$ in this experiment; hence, the adaptive mesh refinement algorithm reduces to a uniform mesh refinement strategy. The computational meshes are depicted in Fig. 2 and the numerical results are presented in Table 1. It can be observed that $\eta/e \approx 1.42$ when $k = 1$, $\eta/e \approx 1.28$ when $k = 2$, and $\eta/e \approx 1.17$ when $k = 3$. This illustrates that our *a posteriori* error estimators are reliable and efficient on uniformly refined meshes.

5.2 *L-shaped domain problem*

Consider the problem (1.1) on an L-shaped domain $\Omega = (-1,1)^2/[0,1] \times [-1,0]$. We take $a = 1$, $f = 0$, $\Gamma_N = \emptyset$, and g_D is chosen corresponding to the following exact solution:

$$u(r,\theta) = r^{2/3}\sin(2\theta/3), \quad \text{and} \quad \boldsymbol{p} = \nabla u,$$

where r, θ are the polar coordinates.

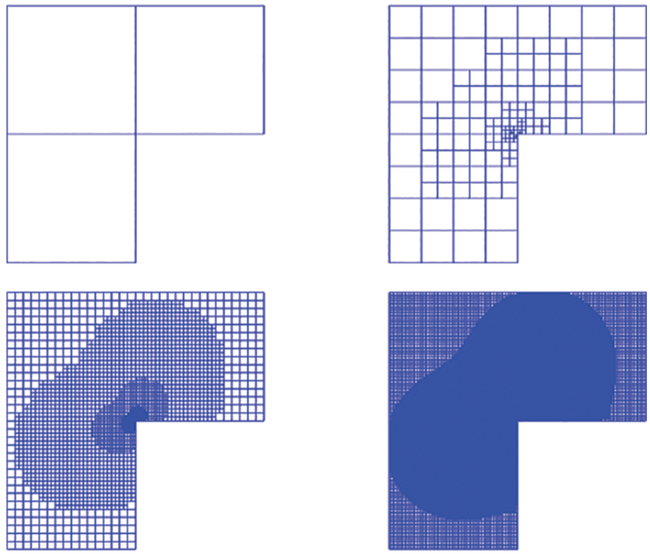

FIG. 3. Adapted meshes for the L-shaped domain problem: levels 0, 6, 12, 18 with $\beta = 0.5$.

TABLE 2 *Results for the L-shaped domain problem*

Mesh	N	e	$N^{-1/2}$	η	$N^{-1/2}/e$	η/e
0	10	4.3233E–01	3.1623E–01	3.0123E–01	0.7314	0.6968
3	85	1.6443E–01	1.0847E–01	2.0516E–01	0.6597	1.2477
6	358	6.7515E–02	5.2852E–02	1.0923E–01	0.7828	1.6179
9	1594	2.9380E–02	2.5047E–02	5.2453E–02	0.8525	1.7853
12	6906	1.3543E–02	1.2033E–02	2.5220E–02	0.8885	1.8622
15	29283	6.4809E–03	5.8438E–03	1.2241E–02	0.9017	1.8888
16	46337	5.0493E–03	4.6455E–03	9.6411E–03	0.9200	1.9094
17	75053	3.9592E–03	3.6502E–03	7.5997E–03	0.9219	1.9195
18	122065	3.1443E–03	2.8622E–03	5.9920E–03	0.9103	1.9057
19	192199	2.4575E–03	2.2810E–03	4.7279E–03	0.9282	1.9239
20	311507	1.9364E–03	1.7917E–03	3.7318E–03	0.9253	1.9272

We use $k = 1$ in this experiment. The meshes generated by the adaptive algorithm are depicted in Fig. 3 with $\beta = 0.5$. The numerical results are presented in Table 2. The adaptive meshes illustrate that the global *a posteriori* error estimator can effectively capture the singularity of the solution. The displayed results confirm that the *a posteriori* error estimators are reliable and efficient.

5.3 *Kellogg's problem*

Consider the problem (1.1) on $\Omega = (0,1)^2$, where the coefficient a is piecewise constant such that $a = a_1$ in the first and third quadrants, and $a = a_2$ in the second and fourth quadrants. We set $\Gamma_N = \emptyset$

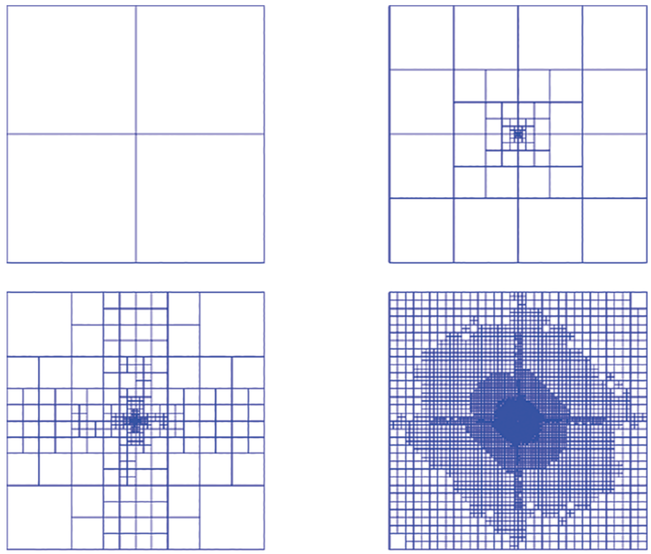

FIG. 4. Adapted meshes for Kellogg's problem: levels 0, 30, 60, 90 with $\beta = 0.3$.

TABLE 3 *Results for Kellogg's problem*

Mesh	N	e	$N^{-1/2}$	η	$N^{-1/2}/e$	η/e
0	12	5.7504E–01	2.8868E–01	1.6179E+00	0.5020	2.8135
10	152	5.0574E–01	8.1111E–02	1.5580E+00	0.1604	3.0807
20	292	4.1897E–01	5.8521E–02	1.3767E+00	0.1397	3.2859
30	440	3.3645E–01	4.7673E–02	1.1553E+00	0.1417	3.4338
40	727	2.4631E–01	3.7088E–02	8.7407E–01	0.1506	3.5486
50	1527	1.6251E–01	2.5591E–02	5.2105E–01	0.1575	3.2062
60	3268	1.1905E–01	1.7493E–02	2.7232E–01	0.1469	2.2874
70	7449	7.8859E–02	1.1586E–02	1.4185E–01	0.1469	1.7987
80	19730	4.4845E–02	7.1193E–03	7.1553E–02	0.1588	1.5955
90	60468	2.3788E–02	4.0667E–03	3.6374E–02	0.1710	1.5291
100	207199	1.2295E–02	2.1969E–03	1.8306E–02	0.1787	1.4889

and $f = 0$. The exact solution in polar coordinates is taken to be $u(r,\theta) = r^{\gamma}\mu(\theta)$, where

$$\mu(\theta) = \begin{cases} \cos[(0.5\pi - \sigma)\gamma]\cos[(\theta - 0.5\pi + \rho)\gamma], & 0 \leqslant \theta \leqslant 0.5\pi, \\ \cos(\rho\gamma)\cos[(\theta - \pi + \sigma)\gamma], & 0.5\pi \leqslant \theta \leqslant \pi, \\ \cos(\sigma\gamma)\cos[(\theta - \pi - \rho)\gamma], & \pi \leqslant \theta \leqslant 1.5\pi, \\ \cos[(0.5\pi - \rho)\gamma]\cos[(\theta - 1.5\pi - \sigma)\gamma], & 1.5\pi \leqslant \theta \leqslant 2\pi, \end{cases} \tag{5.1}$$

and the constants are given by $\gamma = 0.1$, $\rho = 0.25\pi$, $\sigma = -4.75\pi$, $a_1 = 161.4476387975881$ and $a_2 = 1$. Note that the exact solution u belongs to $H^{1+\gamma}(\Omega)$ (see Kellogg, 1975).

We use $k = 1$ in this experiment. The meshes generated by the adaptive algorithm with $\beta = 0.3$ are depicted in Fig. 4. The numerical results are presented in Table 3. The adaptive meshes illustrate that the global *a posteriori* error estimator can effectively capture the singularity of the solution. The displayed error results clearly verify the theoretical results.

Acknowledgements

The authors would like to thank the editors and two referees for their valuable comments and suggestions. In particular, with their help, we managed to remove the quasi-monotonicity assumption on the diffusion coefficient and simplify the error analysis. The authors would also thank Dr John Singler for his help on editing, which significantly improved this paper.

Funding

National Natural Science Foundation of China (NSFC) [grant no. 11801063 to G.C. and grant no. 11771367 to J.C.], China Postdoctoral Science Foundation [grant no. 2018M633339 to G.C.], Key Laboratory of Numerical Simulation of Sichuan Province (Neijiang, Sichuan) [grant no. 2017KF003 to G.C.] and Hong Kong RGC General Research Fund (GRF) [grant no. 15302518 to J.C.].

References

Adams, R. A. & Fournier, J. J. F. (2003) *Sobolev Spaces*. Pure and Applied Mathematics (Amsterdam), vol. 140, 2nd edn. Amsterdam: Elsevier/Academic Press, pp. xiv+305.

Arnold, D. N. (1982) An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, **19**, 742–760.

Arnold, D. N., Brezzi, F., Cockburn, B. & Marini, L. D. (2002) Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, **39**, 1749–1779.

Brenner, S. C. & Scott, L. R. (2008) *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics, vol. 15, 3rd edn. New York: Springer, pp. xviii+397.

Cai, Z., He, C. & Zhang, S. (2017) Discontinuous finite element methods for interface problems: robust a priori and a posteriori error estimates. *SIAM J. Numer. Anal.*, **55**, 400–418.

Cai, Z., Ye, X. & Zhang, S. (2011) Discontinuous Galerkin finite element methods for interface problems: a priori and a posteriori error estimations. *SIAM J. Numer. Anal.*, **49**, 1761–1787.

Carrero, J., Cockburn, B. & Schötzau, D. (2006) Hybridized globally divergence-free LDG methods. I. The Stokes problem. *Math. Comp.*, **75**, 533–563 (electronic).

Castillo, P., Cockburn, B., Perugia, I. & Schötzau, D. (2000) An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, **38**, 1676–1706 (electronic).

Chen, H., Qiu, W., Shi, K. & Solano, M. (2017) A superconvergent HDG method for the Maxwell equations. *J. Sci. Comput.*, **70**, 1010–1029.

Cockburn, B., Gopalakrishnan, J. & Lazarov, R. (2009) Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second-order elliptic problems. *SIAM J. Numer. Anal.*, **47**, 1319–1365.

Cockburn, B., Gopalakrishnan, J. & Sayas, F.-J. (2010) A projection-based error analysis of HDG methods. *Math. Comp.*, **79**, 1351–1367.

Cockburn, B., Kanschat, G. & Schötzau, D. (2005) A locally conservative LDG method for the incompressible Navier-Stokes equations. *Math. Comp.*, **74**, 1067–1095 (electronic).

Cockburn, B. & Shu, C.-W. (1998) The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, **35**, 2440–2463 (electronic).

Cockburn, B. & Zhang, W. (2012) A posteriori error estimates for HDG methods. *J. Sci. Comput.*, **51**, 582–607.

Cockburn, B. & Zhang, W. (2013) A posteriori error analysis for hybridizable discontinuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, **51**, 676–693.

Demkowicz, L. & Gopalakrishnan, J. (2011) Analysis of the DPG method for the Poisson equation. *SIAM J. Numer. Anal.*, **49**, 1788–1809.

Di Pietro, D. A. & Ern, A. (2012) Analysis of a discontinuous Galerkin method for heterogeneous diffusion problems with low-regularity solutions. *Numer. Methods Partial Differ. Equ.*, **28**, 1161–1177.

Dryja, M., Sarkis, M. V. & Widlund, O. B. (1996) Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, **72**, 313–348.

Falgout, R. D. & Yang, U. M. (2002) *Hypre: A Library of High Performance Preconditioners*. Berlin, Heidelberg: Springer, pp. 632–641.

Grisvard, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Monographs and Studies in Mathematics, vol. 24. Boston, MA: Pitman (Advanced Publishing Program), pp. xiv+410.

Gudi, T. (2010) A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.*, **79**, 2169–2189.

Kellogg, R. B. (1975) On the Poisson equation with intersecting interfaces. *Applicable Anal.*, **4**, 101–129. Collection of articles dedicated to Nikolai Ivanovich Muskhelishvili.

Lehrenfeld, C. (2010) Hybrid discontinuous Galerkin methods for solving incompressible flow problems. *Ph.D. Thesis, Rheinisch-Westfalischen Technischen Hochschule Aachen*.

Li, B. & Xie, X. (2016) Analysis of a family of HDG methods for second-order elliptic problems. *J. Comput. Appl. Math.*, **307**, 37–51.

Qiu, W., Shen, J. & Shi, K. (2018) An HDG method for linear elasticity with strong symmetric stresses. *Math. Comp.*, **87**, 69–93.

Qiu, W. & Shi, K. (2016a) An HDG method for convection diffusion equation. *J. Sci. Comput.*, **66**, 346–357.

Qiu, W. & Shi, K. (2016b) A superconvergent HDG method for the incompressible Navier-Stokes equations on general polyhedral meshes. *IMA J. Numer. Anal.*, **36**, 1943–1967.

Verfürth, R. (1994) A posteriori error estimation and adaptive mesh-refinement techniques. *Proceedings of the Fifth International Congress on Computational and Applied Mathematics (Leuven, 1992)*, vol. 50., pp. 67–83.

Wihler, T. P. & Rivière, B. (2011) Discontinuous Galerkin methods for second-order elliptic PDE with low-regularity solutions. *J. Sci. Comput.*, **46**, 151–165.