Numerical Algorithms
https://doi.org/10.1007/s11075-022-01300-3

ORIGINAL PAPER

A hybridizable discontinuous triangular spectral element method on unstructured meshes and its *hp*-error estimates

Bingzhen Zhou[1] · Bo Wang[1] · Li-Lian Wang[2] · Ziqing Xie[1]

Received: 18 August 2021 / Accepted: 10 March 2022

Abstract

In this paper, a hybridizable discontinuous triangular spectral element method (HDTSEM) using tensorial nodal basis functions on unstructured meshes is proposed and analyzed. The elemental local basis is constructed from the one-to-one rectangle-to-triangle transform (Li et al., *Lecture Notes in Computational Sciences and Engineering* 76:237–246, 2011) and glued together under the hybridizable discontinuous Galerkin (HDG) framework. This offers much flexibility allowing for mismatch in nodal points across elements, substantial reduction in global degree of freedoms (DoFs) and excellent mesh adaptivity without sacrificing the high accuracy of a typical spectral element method (SEM). Here, optimal L^2-error estimates are obtained on quasi-uniform unstructured meshes and ample numerical results are provided to validate the theoretical results.

Keywords Spectral element method · Hybridizable discontinuous Galerkin method · Unstructured triangular mesh · hp error analysis

✉ Bo Wang
bowang@hunnu.edu.cn

Bingzhen Zhou
zbzhen@smail.hunnu.edu.cn

Li-Lian Wang
lilian@ntu.edu.sg

Ziqing Xie
ziqingxie@hunnu.edu.cn

[1] Key Laboratory of Computing and Stochastic Mathematics, Ministry of Education (LCSM), College of Mathematics and Computer Science, Hunan Normal University, Changsha 410081, People's Republic of China

[2] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371, Singapore

Published online: 12 April 2022

Mathematics Subject Classification (2010) 65G20 · 65N30 · 65N35 · 65D15

1 Introduction

The spectral element method was first introduced by Patera [29] and had many further developments (merged with the *hp* finite element method) and wide applications in, e.g., fluid dynamics (cf. [2, 9, 20, 23]). The SEM integrates the high-order accuracy of a spectral method with the geometric flexibility and high parallelizability of a finite element method. The standard SEM uses tensorial nodal basis functions on quadrilateral (2D) or hexahedral (3D) meshes that substantially facilitates both the implementation (e.g., the imposition of continuity across elements) and analysis, as many numerical tools and analysis arguments in one dimension can be directly transplanted to multiple dimensions.

In many applications, it is more desirable to employ the SEM on unstructured triangular or tetrahedral meshes (which can be generated automatically by software and produce better approximations to complex domains). In the past two decades, much progress has been made in developing triangular/tetrahedral spectral element method (TSEM). The first is based on the approximation by a nodal basis on special nodal points [4, 16, 17, 28, 36]. The second is the modal TSEM based on Koornwinder-Dubiner (KD) polynomials (cf. [11, 22, 33]). The third is the use of the non-polynomial basis functions constructed by the collapsed Duffy's transform [3, 11, 12, 20, 24, 34, 35] or its important variants [25, 26, 32, 37].

The development of the TSEM using polynomial basis functions has suffered from at least the following three difficulties. Firstly, the optimal nodal set for the construction of nodal basis on triangular/tetrahedral element is difficult to achieve. For example, the well-known Fekete points on triangle is difficult to be generalized to the 3-dimensional case. Secondly, the high-order modal basis based on KD polynomials on simplex is too complicate to be used in the real applications. Thirdly, the TSEM using polynomial basis is less efficient in 3D simulations due to the lack of tensorial structure in the employed high-order basis functions.

As a result, much effort has been devoted to developing TSEMs using the non-polynomial tensorial basis in triangle/tetrahedron [3, 24, 26, 34, 37]. By using some rectangle-to-triangle transforms, e.g., Duffy's transform and the one-to-one transform first introduced in [26], these methods generate rational or irrational basis functions on triangles from standard tensorial polynomial basis functions in reference rectangle. Coordinate transform from a reference cube to a tetrahedron was also suggested in [26], though implementation for 3D problems using tetrahedral meshes has not been reported in the literature so far. The main difficulties in developing the 3D version of these methods are two folds: (i) the singularity induced by the transform will be much stronger; and (ii) the spectral element nodes generated by mapping the tensorial nodes in the reference cube do not match across the elemental interfaces.

In this paper, we propose a hybridizable discontinuous triangular spectral element method (HDTSEM) to overcome the abovementioned difficulties without sacrificing the spectral accuracy. The HDTSEM is drawn on a HDG formulation [6] using

rectangle-to-triangle transform based non-polynomial nodal basis. The HDG formulation has three important advantages. Firstly, the singularity induced by the transform can be canceled by the Jacobian in the weak form of the first-order system adopted in the HDG framework. Secondly, the HDG framework allows the unmatched spectral element nodes. Thirdly, the reduction of the global degree of freedom improves the efficiency significantly when high-order polynomials are used in the construction of non-polynomial basis functions. Moreover, we can rigourously show that the HDTSEM can achieve an optimal convergence rate in terms of both the mesh size h and the polynomial degree p. We provide ample numerical results to further demonstrate the optimality.

The rest of the paper is organized as follows. In Section 2, we start with the rectangle-to-triangle mapping, and then construct the discontinuous spectral element space on unstructured triangular meshes. We demonstrate that the unmatched nodes across the elemental interfaces are unavoidable when we extend the idea to 3D tetrahedral mesh. Then, we formulate and implement the HDTSEM for elliptic problems in Section 3, and conduct rigorous hp-error estimates in Section 4. In Section 5, numerical results are provided to show the flexibility and robustness of the method and also validate the theoretical analysis. Some concluding remarks are given in Section 6.

2 Discontinuous triangular/tetrahedral spectral element spaces

In this section, we construct the discontinuous triangular spectral element approximation spaces based on the rectangle-to-triangle transform [26] of the standard polynomial basis, and discuss the 3D counterpart as well. Due to the singularity of the transform, the resulting spectral element grids usually have unmatched nodes across elemental interfaces. In particular, we can show this situation is unavoidable in the 3D case. To this end, the use of the discontinuous Galerkin technique to glue the elements is deemed desirable.

2.1 The rectangle-to-triangle transform

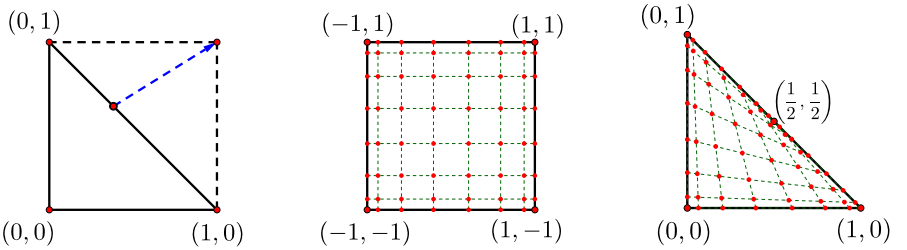
The collapsed Duffy's transform adopted in [11, 12, 20, 24, 34, 35] can be generalized to the rectangle-to-triangle transform which pulls one edge of the triangle into two edges of the rectangle at a given point on this edge (cf. [26, 37], see Fig.1 (a)). In particular, one can choose the middle point resulting the transform (cf. [26, 32, 37]):

$$T(\xi,\eta): \quad \hat{x}=\frac{1}{8}(1+\xi)(3-\eta), \quad \hat{y}=\frac{1}{8}(3-\xi)(1+\eta), \quad \forall(\xi,\eta)\in\square, \tag{2.1}$$

where

$$\triangle := \{(\hat{x},\hat{y}) : 0<\hat{x},\hat{y}<1,\ \ 0<\hat{x}+\hat{y}<1\}, \quad \square := \{(\xi,\eta) : -1<\xi,\eta<1\},$$

are the standard *reference triangle* and the *reference square*, respectively. The above transform T pulls the hypotenuse of $\triangle$ into two edges of $\square$ at the point $(\frac{1}{2},\frac{1}{2})$ (see Fig. 1).



(a) $T : \square \mapsto \triangle$ (b) LGL points on $\square$ (c) mapped LGL points on $\triangle$

Fig. 1 An illustration of the transform $T : \square \to \triangle$ defined in (2.1). (a) $T : \square \to \triangle$. (b) LGL points on $\square$. (c) Mapped LGL points on $\triangle$

The Jacobian matrix and determinant of (2.1) are

$$\mathbb{J}(\xi,\eta) = \begin{bmatrix} \frac{\partial \hat{x}}{\partial \xi} & \frac{\partial \hat{y}}{\partial \xi} \\ \frac{\partial \hat{x}}{\partial \eta} & \frac{\partial \hat{y}}{\partial \eta} \end{bmatrix} = \begin{bmatrix} \frac{3-\eta}{8} & -\frac{1+\eta}{8} \\ -\frac{1+\xi}{8} & \frac{3-\xi}{8} \end{bmatrix}, \tag{2.2}$$

and

$$J(\xi,\eta) = \det(\mathbb{J}) = \frac{2-\xi-\eta}{16}. \tag{2.3}$$

Alternatively, the Jacobian in $(\hat{x},\hat{y})$-coordinates is given by

$$J_{\triangle}(\hat{x},\hat{y}) = \frac{\sqrt{(\hat{x}-\hat{y})^2 + 4(1-\hat{x}-\hat{y})}}{8}. \tag{2.4}$$

Throughout this paper, the subscript $\triangle$ is used to indicate that the arguments are in the reference triangular coordinates $(\hat{x},\hat{y})$. Moreover, the inverse transform is given by

$$T^{-1}(\hat{x},\hat{y}) : \begin{cases} \xi = 1+\hat{x}-\hat{y}-\sqrt{(\hat{x}-\hat{y})^2+4(1-\hat{x}-\hat{y})}, \\ \eta = 1-\hat{x}+\hat{y}-\sqrt{(\hat{x}-\hat{y})^2+4(1-\hat{x}-\hat{y})}, \end{cases} \tag{2.5}$$

with the Jacobian matrix

$$\mathbb{J}^{-1}(\xi,\eta) = \begin{bmatrix} \frac{\partial \xi}{\partial \hat{x}} & \frac{\partial \eta}{\partial \hat{x}} \\ \frac{\partial \xi}{\partial \hat{y}} & \frac{\partial \eta}{\partial \hat{y}} \end{bmatrix} = \frac{1}{8J(\xi,\eta)} \begin{bmatrix} 3-\xi & 1+\eta \\ 1+\xi & 3-\eta \end{bmatrix}. \tag{2.6}$$

If the transform is confined on the elemental edges, we have

$$\begin{aligned} &T(\xi,1): \ \hat{x} = \tfrac{1}{4}(1+\xi), \quad \hat{y} = \tfrac{1}{4}(3-\xi); \quad T(\xi,-1): \ \hat{x} = \tfrac{1}{2}(1+\xi), \quad \hat{y} = 0; \\ &T(1,\eta): \ \hat{x} = \tfrac{1}{4}(3-\eta), \quad \hat{y} = \tfrac{1}{4}(1+\eta); \quad T(-1,\eta): \ \hat{x} = 0, \quad \hat{y} = \tfrac{1}{2}(1+\eta). \end{aligned} \tag{2.7}$$

Apparently, $T : \square \to \triangle$ becomes linear in such cases.

Compared with the Duffy's transform, (2.1) defines a one-to-one mapping with a much weaker singularity [32]. Actually, $1/J(\xi,\eta)$ is integrable as

$$\int_{\square} \frac{1}{2-\xi-\eta} d\xi d\eta = 4\ln 2, \tag{2.8}$$

while that of the Duffy's transform is not integrable.

2.2 Discontinuous triangular spectral element spaces

Hereafter, let $I=(-1,1)$, and for any integer $p\geq 1$, denote by $\mathcal{P}_p(I)$ the set of all algebraic polynomials of degree at most p. Two standard polynomial spaces on the reference $\triangle$ and $\square$ are

$$\mathcal{P}_p(\triangle) := \mathrm{span}\{\hat{x}^i\hat{y}^j \,:\, 0\leq i+j\leq p\}, \quad \mathcal{Q}_p(\square) := \mathcal{P}_p(I)\otimes\mathcal{P}_p(I). \tag{2.9}$$

Define the non-polynomial approximation space

$$Y_p(\triangle) = \mathcal{Q}_p(\square)\circ T^{-1} = (\mathcal{P}_p(I)\otimes\mathcal{P}_p(I))\circ T^{-1}, \tag{2.10}$$

which consists of all images of the tensorial polynomials in $\mathcal{Q}_p(\square)$ under the inverse transform T^{-1} in (2.5). In practice, we use the nodal basis of $\mathcal{Q}_p(\square)$. Denote by $\{\zeta_i\}_{i=0}^p$ the Legendre-Gauss-Lobatto (LGL) points on $\bar{I}$ and $\{h_m(\zeta)\}_{m=0}^p$ the corresponding Lagrange interpolation basis, i.e., $h_m\in\mathcal{P}_p(I)$ and $h_m(\zeta_n)=\delta_{mn}$ (where δ_{mn} is the Kronecker Delta symbol). Then

$$\mathcal{Q}_p(\square) = \mathrm{span}\{\varphi_{mn} : \varphi_{mn}(\xi,\eta)=h_m(\xi)h_n(\eta),\ \ 0\leq m,n\leq p\}, \tag{2.11}$$

which leads to the nodal basis of $Y_p(\triangle)$:

$$Y_p(\triangle) = \mathrm{span}\{\psi_{mn} : \psi_{mn}(\hat{x},\hat{y})=\varphi_{mn}\circ T^{-1}(\hat{x},\hat{y}),\ \ 0\leq m,n\leq p\}. \tag{2.12}$$

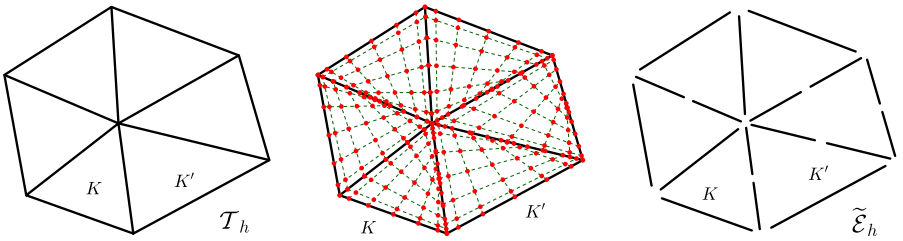
Let $\mathcal{T}_h := \{K\}$ be a shape regular quasi-uniform triangular mesh of a polygonal domain Ω in the sense that there exist positive constants c_0 and c_1 such that

$$\max_{K\in\mathcal{T}_h}\frac{h_K^2}{|K|}\leq c_0, \quad \frac{h}{\min\limits_{K\in\mathcal{T}_h} h_K}\leq c_1. \tag{2.13}$$

Here, we denote by $h_K=\mathrm{diam}(K)$ the diameter of element K, $|K|$ the measure of K, mesh size $h=\max\{h_K\}$. On each element K, we define the local spectral element spaces

$$V_h^p(K) = \left\{v_h^p \,:\, v_h^p\circ\mathscr{F}_K\in Y_p(\triangle)\right\}, \quad \boldsymbol{V}_h^p(K) = \left(V_h^p(K)\right)^2, \tag{2.14}$$

where $\mathscr{F}_K:\triangle\to K$ the standard affine mapping from the reference triangle $\triangle$ to a physical element $K\in\mathcal{T}_h$. It is evident that the nodal basis of the above local spectral element spaces is associated with the spectral element nodes which are the images of the LGL points under the mapping $\mathscr{F}_K\circ T$ (see Fig. 2 (c)). Note that the spectral element grid may not match along the interior edges, for example, along the interior edge $K\cap K'$ in Fig. 2 (b). Although the inconsistency can be avoided in the 2D case with a careful orientation of the singular edges in the unstructured meshes (cf. [37]), it becomes not possible for the 3D tetrahedral mesh (see some detailed discussions in the next subsection). As a result, we feel compelled to develop the discontinuous



(a) triangular mesh (b) spectral element grids with mismatched nodes (c) $\widetilde{\mathcal{E}}_h$: refined mesh skeleton

Fig. 2 Spectral element grid generation on a simple triangular mesh with polynomial degree $p = 6$. (a) Triangular mesh. (b) Spectral element grids with mismatched nodes. (c) $\widetilde{\mathcal{E}}_h$: refined mesh skeleton

Galerkin method to achieve spectral accuracy in the presence of unmatched nodes. For this purpose, we define the discontinuous triangular spectral element spaces on $\mathcal{T}_h$ as follows

$$V_h^p(\mathcal{T}_h) = \left\{ v_h^p \in L^2(\Omega) \,:\, v_h^p|_K \in V_h^p(K), \forall K \in \mathcal{T}_h \right\}, \quad \boldsymbol{V}_h^p(\mathcal{T}_h) = \left(V_h^p(\mathcal{T}_h)\right)^2. \quad (2.15)$$

2.3 Discontinuous tetrahedral spectral element spaces

A similar one-to-one transform from the reference cube $\square_3 := (-1, 1)^3$ to the reference tetrahedron $\triangle_3 = \{(\hat{x}, \hat{y}, \hat{z}) : 0 < \hat{x}, \hat{y}, \hat{z}, \hat{x} + \hat{y} + \hat{z} < 1\}$ is also available (cf. [26]):

$$T_3(\xi, \eta, \zeta) : \begin{cases} \hat{x} = \dfrac{1}{24}(1+\xi)(7-2\eta-2\zeta+\eta\zeta), \\ \hat{y} = \dfrac{1}{24}(1+\eta)(7-2\xi-2\zeta+\xi\zeta), \\ \hat{z} = \dfrac{1}{24}(1+\zeta)(7-2\xi-2\eta+\xi\eta). \end{cases} \quad (2.16)$$

This transform pulls the center of the face with vertices $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ and the middle points of its three edges to vertices of the reference cube $\square_3$ (see Fig. 3). Define the tensorial polynomial space

$$\mathcal{Q}_p(\square_3) = \text{span}\{\varphi_{ijk} : \varphi_{ijk}(\xi, \eta, \zeta) = h_i(\xi)h_j(\eta)h_k(\zeta), \quad 0 \le i, j, k \le p\}. \quad (2.17)$$

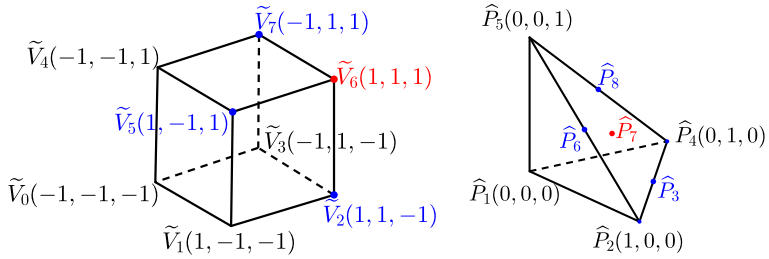
Then the nodal basis of $Y_p(\triangle_3) := \mathcal{Q}_p(\square_3) \circ T_3^{-1}$ is given by

$$Y_p(\triangle_3) = \text{span}\{\psi_{ijk} : \psi_{ijk}(\hat{x}, \hat{y}, \hat{z}) = \varphi_{ijk} \circ T_3^{-1}(\hat{x}, \hat{y}, \hat{z}), \quad 0 \le i, j, k \le p\}. \quad (2.18)$$

Let $\mathcal{T}_h^{3d} = \{K\}$ be a shape regular quasi-uniform tetrahedral mesh on a polyhedral region Ω. On each tetrahedral element K, we define the local spectral element spaces

$$V_h^p(K) = \left\{ v_h^p : v_h^p \circ \mathscr{F}_K \in Y_p(\triangle_3) \right\}, \quad \boldsymbol{V}_h^p(K) = \left(V_h^p(K)\right)^3, \quad (2.19)$$

where $\mathscr{F}_K$ is the standard affine mapping from the reference tetrahedron $\triangle_3$ to the physical tetrahedron K. Apparently, the nodes associated with the above local spectral element spaces are unmatched on some interior faces and edges. Below, we will



(a) Reference cube (b) Reference tetrahedron

Fig. 3 Reference cube $\square_3 = (-1,1)^3$ and reference tetrahedron $\triangle_3$

give an example to show that the unmatched nodes always occur even for a very simple tetrahedral mesh.

By tuning the affine mapping $\mathscr{F}_K : \triangle_3 \to K$, one can determine which face of the tetrahedron K will be pulled into three faces in the reference cube under the transform $T_3^{-1} \circ \mathscr{F}_K$. For a set of affine mapping $\{\mathscr{F}_K\}$ associated to the elements of $\mathcal{T}_h^{3d}$, we mark the pulled face and its three edges element-by-element. The tetrahedral mesh $\mathcal{T}_h^{3d} = \{K\}$ together with the mapping set $\{\mathscr{F}_K\}$ generates matched spectral element grids if and only if each element $K \in \mathcal{T}_h^{3d}$ has and only has one face and its edges marked (see the two situations in Fig. 4 (a)–(b)).

Suppose a tetrahedral mesh $\mathcal{T}_h^{3d}$ together with a set of affine mappings $\{\mathscr{F}_K\}$ and the cube-to-tetrahedron mapping T_3 generate matched spectral element nodes. Then, regarding the marked faces and edges in $\mathcal{T}_h^{3d} = \{K\}$ determined by the mapping set $\{\mathscr{F}_K\}_{K\in\mathcal{T}_h}$, we have the following two observations:

(i) For two adjacent elements $K = \{V_i, V_j, V_k, V_l\}$ and $K' = \{V_i, V_j, V_k, V_m\}$ in $\mathcal{T}_h^{3d}$, if edge V_iV_j is marked, then either the shared face $\triangle_{V_iV_jV_k}$ is marked (see Fig. 4 (a)), or both faces $\triangle_{V_iV_jV_l}$ and $\triangle_{V_iV_jV_m}$ are marked (see Fig. 4 (b)).

(ii) If $K_1 = \{V_i, V_j, V_l, V_k\}$, $K_2 = \{V_i, V_j, V_m, V_k\}$ and $K_3 = \{V_i, V_j, V_l, V_m\}$ are three adjacent elements such that $K_1 \cap K_2 \cap K_3 = V_iV_j$, and the edge

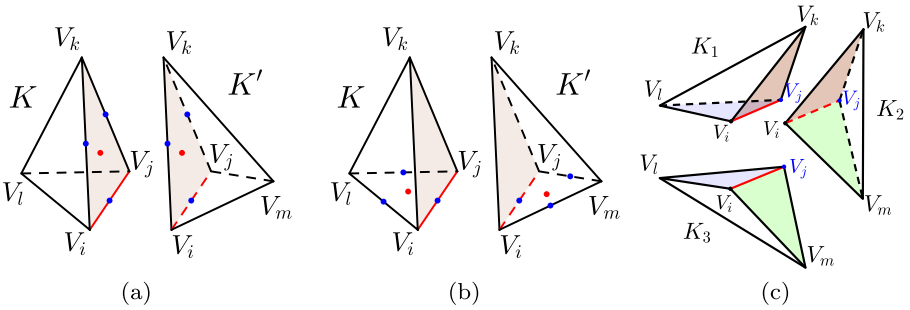


Fig. 4 (a) Two adjacent elements with a shared face marked; (b) two adjacent elements with a shared face not marked; (c) an interior edge shared by three elements

$V_iV_j \notin K$, for $\forall K \in \mathcal{T}_h^{3d}/\{K_1, K_2, K_3\}$, $V_iV_j \notin \partial\Omega^{3d}$, then the edge V_iV_j and the faces $\triangle_{V_iV_iV_k}$, $\triangle_{V_iV_iV_l}$ and $\triangle_{V_iV_iV_m}$ are not marked (see Fig. 4 (c)).

The first observation is straightforward and the second one can be justified by contradiction. Actually, if edge V_iV_j is marked, then one of the three faces $\triangle_{V_iV_jV_k}$, $\triangle_{V_iV_jV_l}$ and $\triangle_{V_iV_jV_m}$ must be marked. Without loss of generality, we assume that the face $\triangle_{V_iV_jV_k}$ is marked. Applying Observation (i) to the adjacent elements K_1, K_3, we conclude that the face $\triangle_{V_iV_jV_m}$ is marked, i.e., the element K_2 has two marked faces. This contradicts the basic fact that each element in a node matched spectral element grid can not have two marked faces.

Now, we are ready to give a simple tetrahedral mesh $\mathcal{T}_c$ (see Fig. 5) which unavoidably generate unmatched nodes under the mapping $T_3 : \square_3 \to \triangle_3$. The six vertices are

$$V_1(0, 0, \sqrt{3}), \quad V_2(0, 0, -\sqrt{3}), \quad V_3(1, \sqrt{3}, 0), \quad V_4(1, -\sqrt{3}, 0),$$
$$V_5(3, 0, 0), \quad V_6(-2, 0, 0),$$

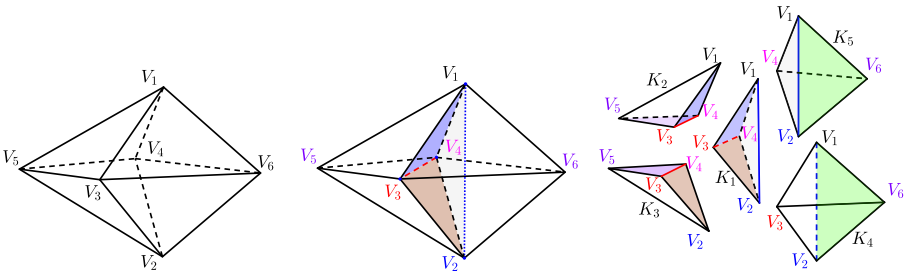
and the five elements are

$$K_1 = \{V_1V_2V_3V_4\}, \quad K_2 = \{V_1V_3V_4V_5\}, \quad K_3 = \{V_2V_3V_4V_5\},$$
$$K_4 = \{V_1V_2V_6V_3\}, \quad K_5 = \{V_1V_2V_6V_4\}$$

in $\mathcal{T}_c$. Assume that there is an affine mapping set $\{\mathcal{F}_K\}$ such that a consistent spectral element grid can be generated for $\mathcal{T}_c$ by using the transform T_3. In view of Observation (ii), the interior edges V_1V_2 and V_3V_4 shared by elements $\{K_1, K_4, K_5\}$, and $\{K_1, K_2, K_3\}$, respectively, are not marked, and the faces $\triangle_{V_1V_2V_3}$, $\triangle_{V_1V_2V_4}$, $\triangle_{V_1V_3V_4}$ and $\triangle_{V_2V_3V_4}$ are not marked. That is the element K_1 does not have any marked face which contradicts the fact that each element must have one marked face.

3 HDTSEM: Implementation

In this section, we introduce the HDTSEM for an elliptic problem, and describe the detailed implementation.



(a) Domain (b) Mesh $\mathcal{T}_c$ (c) Element splitting

Fig. 5 A simple tetrahedral mesh which can not generate a consistent spectral element grid under the mapping $T_3 : \square_3 \to \triangle_3$. (a) Domain. (b) Mesh T_c. (c) Element splitting

3.1 The HDG scheme

Consider the elliptic boundary value problem

$$-\nabla\cdot(\beta(\boldsymbol{x})\nabla u)+\gamma(\boldsymbol{x})u=f(\boldsymbol{x})\quad\text{in }\Omega;\quad u|_{\partial\Omega}=0, \tag{3.20}$$

where $\boldsymbol{x}=(x,y)$ and Ω is an open, bounded and polygonal domain. Assume that the coefficients β,γ are given positive functions in Ω. By introducing an auxiliary variable $\boldsymbol{q}$, the elliptic (3.20) can be rewritten as the following first-order system

$$\begin{aligned}\boldsymbol{q}+\beta\nabla u&=0\quad\text{in }\Omega,\\ \nabla\cdot\boldsymbol{q}+\gamma u&=f\quad\text{in }\Omega,\\ u&=0\quad\text{on }\partial\Omega.\end{aligned} \tag{3.21}$$

The HDG method introduces an approximation for the unknown function u on the mesh skeleton $\mathcal{E}_h:=\{e:e\in\partial K,\forall K\in\mathcal{T}_h\}$. Therefore, a piecewise polynomial space defined on the mesh skeleton is needed. We call the edge $e\in K$ is pulled edge of the triangular element $K\in\mathcal{T}_h$, if the middle point of edge e is the image of the vertex $(1,1)$ under the mapping $\mathscr{F}_K\circ T$, where $\mathscr{F}_K$ is the standard affine mapping from the reference triangle $\triangle$ to K. Noting that the HDG framework will be adopted, the nodes in neighboring elements are allowed to be mismatched on the shared edge, see Fig. 2 (b) for an illustration. Therefore, we can set the pulled edge for each element freely. This is the main advantage of using DG framework compared with our previous work [37] using CG framework. The mesh skeleton $\mathcal{E}_h$ will be decomposed into two parts $\mathcal{E}_h=\mathcal{E}_h^{(1)}\cup\mathcal{E}_h^{(2)}$, where

$$\begin{aligned}\mathcal{E}_h^{(2)}&:=\{e:e=K\cap K'\in\mathcal{E}_h,\ \text{is a pulled edge of element } K \text{ or } K'\},\\ \mathcal{E}_h^{(1)}&:=\mathcal{E}_h\backslash\mathcal{E}_h^{(2)}.\end{aligned}$$

As depicted in Fig. 2 (b), the interior edge $e=K\cap K'\in\mathcal{E}_h^{(2)}$, is the pulled edge of the element K, but not the pulled edge of the element K'. For any edge $e\in\mathcal{E}_h^{(2)}$, we divide it into two subedges e_1 and e_2 from its middle point. Denote by $\widetilde{\mathcal{E}}_h^{(2)}$ the set of all subedges and define a refined mesh skeleton as follows (see Fig. 2 (c)):

$$\widetilde{\mathcal{E}}_h=\mathcal{E}_h^{(1)}\cup\widetilde{\mathcal{E}}_h^{(2)}. \tag{3.22}$$

Accordingly, the approximate trace space on $\widetilde{\mathcal{E}}_h$ is defined as

$$\mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)=\left\{\mu_h^p\in L^2(\widetilde{\mathcal{E}}_h):\mu_h^p|_e\circ\mathscr{F}_e\in\mathcal{P}_p(I),\forall e\in\widetilde{\mathcal{E}}_h\right\}, \tag{3.23}$$

where $\mathscr{F}_e$ is the linear transform from the reference interval $I:=(-1,1)$ to an edge $e\in\widetilde{\mathcal{E}}_h$.

The HDTSEM is to find $(\boldsymbol{q}_h^p,u_h^p,\hat{u}_h^p)\in\boldsymbol{V_h^p}(\mathcal{T}_h)\times V_h^p(\mathcal{T}_h)\times\mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$ such that

$$(\beta^{-1}\boldsymbol{q}_h^p,\boldsymbol{v})_{\mathcal{T}_h}-(u_h^p,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h}+\langle\hat{u}_h^p,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}=0, \tag{3.24a}$$

$$-(\boldsymbol{q}_h^p,\nabla w)_{\mathcal{T}_h}+(\gamma u_h^p,w)_{\mathcal{T}_h}+\langle\hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n},w\rangle_{\partial\mathcal{T}_h}=(f,w)_{\mathcal{T}_h}, \tag{3.24b}$$

$$\langle\hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega}=0, \tag{3.24c}$$

$$\langle\hat{u}_h^p,\mu\rangle_{\partial\Omega}=0, \tag{3.24d}$$

for all $(\boldsymbol{v}, w, \mu) \in \boldsymbol{V_h^p}(\mathcal{T}_h) \times V_h^p(\mathcal{T}_h) \times \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$, and on $\partial\mathcal{T}_h$, we require

$$\hat{\boldsymbol{q}}_h^p = \boldsymbol{q}_h^p + \tau \frac{p}{h}(u_h^p - \hat{u}_h^p)\boldsymbol{n}, \tag{3.25}$$

with $\boldsymbol{n}$ is the outward unit normal vector of $\partial\mathcal{T}_h$, and τ is a positive and mesh independent constant (cf. [31]). In (3.24a) and what follows, we denote

$$(\boldsymbol{u}, \boldsymbol{v})_{\mathcal{T}_h} := \sum_{i=1}^{d} (u_i, v_i)_{\mathcal{T}_h}, \quad (u, v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (u, v)_K, \quad \langle u, v\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle u, v\rangle_{\partial K},$$

where $(u, v)_K$ and $\langle u, v\rangle_{\partial K}$ denote the integral of uv over an element K and its edge ∂K, respectively.

3.2 Efficient implementation

The HDTSEM solutions on any given triangle $K_j \in \mathcal{T}_h$ are given by

$$\boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}|_{K_j} = \sum_{m=0}^{p}\sum_{n=0}^{p} \boldsymbol{Q}_{mn}^{(j)} \psi_{mn}(\mathscr{F}_{K_j}^{-1}(x, y)), \; u_h^p|_{K_j} = \sum_{m=0}^{p}\sum_{n=0}^{p} U_{mn}^{(j)} \psi_{mn}(\mathscr{F}_{K_j}^{-1}(x, y)), \tag{3.26}$$

while the approximate trace can be represented as

$$\hat{u}_h^p|_{F_i} = \sum_{n=0}^{p} \widehat{U}_n^{(i)} h_n(\mathscr{F}_{F_i}^{-1}(x, y)), \quad \forall F_i \in \mathcal{E}_h. \tag{3.27}$$

Substituting (3.26)–(3.27) into (3.24a)–(3.25) leads to the linear system

$$\begin{bmatrix} \mathbb{K}_{11} & \mathbb{K}_{12} & \mathbb{K}_{13} \\ \mathbb{K}_{21} & \mathbb{K}_{22} & \mathbb{K}_{23} \\ \mathbb{K}_{31} & \mathbb{K}_{32} & \mathbb{K}_{33} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \\ \widehat{\boldsymbol{U}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{F} \\ \boldsymbol{0} \end{bmatrix}, \tag{3.28}$$

where the global matrices $\mathbb{K}_{nm}(n, m = 1, 2, 3)$ and the right-hand side $\boldsymbol{F}$ is obtained from the local contributions (or elements) $\mathbb{K}_{nm}^{(j)}(n, m = 1, 2, 3)$, $\boldsymbol{F}^{(j)}$ via sub-assembly. The unknown vectors $\boldsymbol{Q}, \boldsymbol{U}$ and $\widehat{\boldsymbol{U}}$ consist of nodal values of $\boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}$, u_h^p and $\hat{u}_h^p$. As in the HDG method, $\boldsymbol{Q}$ and $\boldsymbol{U}$ are local unknowns which can be solved from $\widehat{\boldsymbol{U}}$ element-by-element, i.e.,

$$\begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \end{bmatrix} = \mathbb{A}^{-1} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{F} \end{bmatrix} - \mathbb{A}^{-1} \begin{bmatrix} \mathbb{K}_{13} \\ \mathbb{K}_{23} \end{bmatrix} \widehat{\boldsymbol{U}}, \quad \mathbb{A} := \begin{bmatrix} \mathbb{K}_{11} & \mathbb{K}_{12} \\ \mathbb{K}_{21} & \mathbb{K}_{22} \end{bmatrix}, \tag{3.29}$$

and $\mathbb{A}$ is block diagonal. Therefore, the linear system (3.28) can be reduced to

$$\left(\begin{bmatrix} \mathbb{K}_{31} & \mathbb{K}_{32} \end{bmatrix} \mathbb{A}^{-1} \begin{bmatrix} \mathbb{K}_{13} \\ \mathbb{K}_{23} \end{bmatrix} - \mathbb{K}_{33} \right) \widehat{\boldsymbol{U}} = \begin{bmatrix} \mathbb{K}_{31} & \mathbb{K}_{32} \end{bmatrix} \mathbb{A}^{-1} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{F} \end{bmatrix}. \tag{3.30}$$

We emphasize that $\widehat{\boldsymbol{U}}$ only involves the unknowns on the mesh skeleton. The global unknowns in (3.30) is of order $\mathcal{O}(ph^{-2})$ while the classical quadrilateral SEM has global degree of freedom $\mathcal{O}(p^2h^{-2})$. Although the hybridization or static condensation technique can be applied to the quadrilateral SEM (cf. [15, 19, 21, 30]), the

new HDTSEM is comparable to the classic SEM on quadrilateral meshes in terms of efficiency.

As in the classic SEM, high-order polynomials are more preferable in the HDTSEM for both efficiency and accuracy. However, the usage of high-order polynomials will result in large dense local systems in (3.29). For these dense linear systems, the tensorial structure in the basis functions is crucial for the construction of fast solvers (cf. [9, 15, 19, 30]). As the basis functions $\{\psi_{mn}(\hat{x}, \hat{y})\}$ are constructed by the tensor products of one-dimensional Lagrange basis functions $\{h_m(\xi), h_n(\eta)\}$, we are able to write the local matrices $\mathbb{K}^{(j)}_{nm}(n, m = 1, 2, 3)$ as the Kronecker products of one-dimensional local matrices. For this purpose, we first assume that $\beta(\boldsymbol{x})$ and $\gamma(\boldsymbol{x})$ are piecewise constants and $\beta(\boldsymbol{x}) = \beta_j, \gamma(\boldsymbol{x}) = \gamma_j, \forall \boldsymbol{x} \in K_j$. Define the matrices $\widetilde{\mathbb{M}} = (\widetilde{M}_{nn'}), \widehat{\mathbb{M}} = (\widehat{M}_{nn'}), \widetilde{\mathbb{C}} = (\widetilde{C}_{nn'}), \widehat{\mathbb{C}} = (\widehat{C}_{nn'})$ with the entries

$$\begin{array}{ll} \widetilde{M}_{nn'} = \int_{-1}^{1} h_n(\xi) h_{n'}(\xi) \mathrm{d}\xi, & \widehat{M}_{nn'} = \int_{-1}^{1} h_n(\xi) h_{n'}(\xi) \xi \mathrm{d}\xi, \\ \widetilde{C}_{nn'} = \int_{-1}^{1} h_n(\xi) h'_{n'}(\xi) \mathrm{d}\xi, & \widehat{C}_{nn'} = \int_{-1}^{1} h_n(\xi) h'_{n'}(\xi) \xi \mathrm{d}\xi, \end{array} \tag{3.31}$$

and denote by

$$\Psi^{(j)}_{mn}(x, y) := \psi_{mn}(F^{-1}_{K_j}(x, y)), \quad m, n = 0, 1, \cdots, p, \tag{3.32}$$

i.e., the restrictions of typical basis functions to a given element K_j. Then

$$\begin{array}{l} \int_{K_j} \Psi^{(j)}_{mn}(x, y) \Psi^{(j)}_{m'n'}(x, y) \mathrm{d}x \mathrm{d}y \\ = 2|K_j| \int_{\square} h_m(\xi) h_n(\eta) h_{m'}(\xi) h_{n'}(\eta) J(\xi, \eta) \mathrm{d}\xi \mathrm{d}\eta \\ = \frac{|K_j|}{8} (2\widetilde{M}_{mm'} \widetilde{M}_{nn'} - \widehat{M}_{mm'} \widetilde{M}_{nn'} - \widetilde{M}_{mm'} \widehat{M}_{nn'}), \end{array} \tag{3.33}$$

where the expression (2.3) of $J(\xi, \eta)$ is used. Similarly,

$$\begin{array}{l} \int_{K_j} \Psi^{(j)}_{mn}(x, y) \nabla \Psi^{(j)}_{m'n'}(x, y) \mathrm{d}x \mathrm{d}y \\ = 2|K_j| \mathbb{J}^{-1}_{K_j} \int_{\square} h_m(\xi) h_n(\eta) J \mathbb{J}^{-1} \begin{bmatrix} h'_{m'}(\xi) h_{n'}(\eta) \\ h_{m'}(\xi) h'_{n'}(\eta) \end{bmatrix} \mathrm{d}\xi \mathrm{d}\eta \\ = \frac{|K_j|}{4} \mathbb{J}^{-1}_{K_j} \begin{bmatrix} (3\widetilde{C}_{mm'} - \widehat{C}_{mm'}) \widetilde{M}_{nn'} + \widetilde{M}_{mm'} (\widehat{C}_{nn'} + \widetilde{C}_{nn'}) \\ (\widehat{C}_{mm'} + \widetilde{C}_{mm'}) \widetilde{M}_{nn'} + \widetilde{M}_{mm'} (3\widetilde{C}_{nn'} - \widehat{C}_{nn'}) \end{bmatrix}, \end{array} \tag{3.34}$$

by the expression (2.3) and (2.2). For the integrals along the edges, we have

$$\begin{array}{l} \int_{\partial K_j} \Psi^{(j)}_{mn}(x, y) \Psi^{(j)}_{m'n'}(x, y) \mathrm{d}s \\ = \frac{|e^1_j|}{4} \left(\delta_{mp} \delta_{m'p} \int_{-1}^{1} h_n(\eta) h_{n'}(\eta) \mathrm{d}\eta + \delta_{np} \delta_{n'p} \int_{-1}^{1} h_m(\xi) h_{m'}(\xi) \mathrm{d}\xi \right) \\ + \frac{|e^2_j| \delta_{m0} \delta_{m'0}}{2} \int_{-1}^{1} h_n(\eta) h_{n'}(\eta) \mathrm{d}\eta + \frac{|e^3_j| \delta_{n0} \delta_{n'0}}{2} \int_{-1}^{1} h_m(\xi) h_{m'}(\xi) \mathrm{d}\xi, \end{array} \tag{3.35}$$

where $\{e^1_j, e^2_j, e^3_j\}$ are the edges of the triangle K_j and δ_{mn} is the Kronecker symbol. With the formulations (3.33)–(3.35) and the numerical scheme (3.24a), the local

contributions (or elements) $\mathbb{K}_{nm}^{(j)}(n,m = 1,2)$ can bes written in terms of the Kronecker products as follows

$$\begin{aligned}
\mathbb{K}_{11}^{(j)} =& \frac{|K_j|}{8\beta_j}\mathbb{I}_2 \otimes \big(2\widetilde{\mathbb{M}}\otimes\widetilde{\mathbb{M}} - \widehat{\mathbb{M}}\otimes\widetilde{\mathbb{M}} - \widetilde{\mathbb{M}}\otimes\widehat{\mathbb{M}}\big),\\
\mathbb{K}_{21}^{(j)} =& \Big[\mathbb{K}_x^{(j)}\ \mathbb{K}_y^{(j)}\Big], \quad \mathbb{K}_{12}^{(j)} = -\big(\mathbb{K}_{21}^{(j)}\big)^{\mathrm{T}},\\
\mathbb{K}_x^{(j)} =& \frac{|K_j|}{4}\big[a_{11}^{(j)}\widetilde{\mathbb{C}} + a_{12}^{(j)}\widehat{\mathbb{C}}\big)\otimes\widetilde{\mathbb{M}} + \widetilde{\mathbb{M}}\otimes\big(a_{13}^{(j)}\widetilde{\mathbb{C}} + a_{14}^{(j)}\widehat{\mathbb{C}}\big)\big],\\
\mathbb{K}_y^{(j)} =& \frac{|K_j|}{4}\big[a_{21}^{(j)}\widetilde{\mathbb{C}} + a_{22}^{(j)}\widehat{\mathbb{C}}\big)\otimes\widetilde{\mathbb{M}} + \widetilde{\mathbb{M}}\otimes\big(a_{23}^{(j)}\widetilde{\mathbb{C}} + a_{24}^{(j)}\widehat{\mathbb{C}}\big)\big],\\
\mathbb{K}_{22}^{(j)} =& \gamma_j\big(2\widetilde{\mathbb{M}}\otimes\widetilde{\mathbb{M}} - \widehat{\mathbb{M}}\otimes\widetilde{\mathbb{M}} - \widetilde{\mathbb{M}}\otimes\widehat{\mathbb{M}}\big)\\
&+\frac{\tau p}{h}\left[\frac{|e_j^1|}{4}\big(\mathbb{E}_{pp}\otimes\widetilde{\mathbb{M}} + \widetilde{\mathbb{M}}\otimes\mathbb{E}_{pp}\big) + \frac{|e_j^2|}{2}\mathbb{E}_{00}\otimes\widetilde{\mathbb{M}} + \frac{|e_j^3|}{2}\widetilde{\mathbb{M}}\otimes\mathbb{E}_{00}\right],
\end{aligned} \tag{3.36}$$

where $\mathbb{I}_n$ is the identity matrix of size n, $\mathbb{E}_{mn} = (E_{rs})$ is the matrix of size $(p+1)\times(p+1)$ with only one nonzero entry $E_{mn} = 1$, and

$$(a_{mn}^{(j)})_{2\times 4} = \mathbb{J}_{K_j}^{-1}\begin{bmatrix} 3 & -1 & 1 & 1\\ 1 & 1 & 3 & -1\end{bmatrix}.$$

The Kronecker product formulations can also be obtained for $\mathbb{K}_{m3}, \mathbb{K}_{3m}, (m = 1,2,3)$ by using integral formulas similar to (3.35). As they only involve one-dimensional integrals, we omit the detailed expressions for simplicity. To extend the idea to general variable coefficients $\beta(\boldsymbol{x})$, $\gamma(\boldsymbol{x})$, the coefficients are firstly approximated by interpolation, we refer to [9] for more details.

4 HDTSEM: *hp*-error estimates

We first introduce some necessary notation. Given an open bounded domain D, the weighted Sobolev space $H_w^r(D)$ with $r > 0$ is defined as in Adams [1], and its norm and semi-norm are denoted by $\|\cdot\|_{r,w,D}$ and $|\cdot|_{r,w,D}$, respectively. In particular, if $r = 0$, we denote the inner product and norm of $L_w^2(D)$ by $(\cdot,\ \cdot)_{w,D}$ and $\|\cdot\|_{w,D}$, respectively. We further define the broken Sobolev space

$$H_w^s(\mathcal{T}_h) := \prod_{K\in\mathcal{T}_h} H_w^s(K) \tag{4.37}$$

with norm $\|\cdot\|_{r,w,\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\|\cdot\|_{r,w,K}$ and semi-norm $|\cdot|_{r,w,\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}|\cdot|_{r,w,K}$. If $r = 0$, we denote the space by $L_w^2(\mathcal{T}_h)$ and its inner product and norm by $(\cdot,\ \cdot)_{w,\mathcal{T}_h}$ and $\|\cdot\|_{w,\mathcal{T}_h}$, respectively. If $w \equiv 1$, we drop it from the notion.

For any function $v(x,y) \in V_h^p(K)$, we always denote by

$$\hat{v}(\hat{x},\hat{y}) := v\circ\mathscr{F}_K(\hat{x},\hat{y}), \quad \tilde{v}(\xi,\eta) := \hat{v}\circ T(\xi,\eta),$$

the transformed functions in the reference triangle and square, respectively. Further, denote by $\mathbb{J}_K$ the Jacobian matrix of the affine mapping $\mathscr{F}_K$, and by J_K its

determinant, which are constant matrices and scalar quantities depending on the element K.

Due to the use of the $\square \to \triangle$ transform and the non-polynomial basis in the physical space, some useful tools for the analysis of classic SEM are not available in this setting. Therefore, the error analysis for the HDTSEM in (3.24a) can not directly follow the general framework for HDG method developed by Cockburn and collaborators in [7, 8, 31]. Below, we have some insights into the differences and connections of the approximation spaces between the classic SEM and the HDTSEM. Denote the pth-order polynomial space on K by

$$\mathcal{P}_p(K) := \text{span}\{x^i y^j \ : \ 0 \le i + j \le p\}.$$

Proposition 1 *For the discontinuous spectral element spaces $\boldsymbol{V}_{\boldsymbol{h}}^{\boldsymbol{p}}(K)$, $V_h^p(K)$, $\mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$ defined in (2.14) and (3.23), we have*

(i) $\mathcal{P}_p(K) \subset V_h^p(K) \subset H^1(K)$, *for all* $K \in \mathcal{T}_h$;
(ii) $(J \circ T \circ \mathscr{F}_K)\nabla v \in \boldsymbol{V}_h^p(K)$ *holds for all* $v \in V_h^p(K), \ K \in \mathcal{T}_h$;
(iii) $V_h^p(K)|_{\partial K} \subset \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)|_{\partial K}$ *holds for all* $K \in \mathcal{T}_h$.

Proof It was shown in [32] that $P_p(\triangle) \subset Y_p(\triangle)$. As $\mathscr{F}_K$ is an affine mapping from $\triangle$ to K, the definition (2.14) of $V_h^p(K)$ implies $\mathcal{P}_p(K) \subset V_h^p(K)$.

Given $v \in V_h^p(K), \forall K \in \mathcal{T}_h$, we have

$$\int_K |\nabla v|^2 \mathrm{d}x\mathrm{d}y = \int_\triangle J_K |\mathbb{J}_K^{-1}\widehat{\nabla}\hat{v}|^2 \mathrm{d}\hat{x}\mathrm{d}\hat{y} = \int_\square J_K J |\mathbb{J}_K^{-1}\mathbb{J}^{-1}\widetilde{\nabla}\tilde{v}|^2 \mathrm{d}\xi\mathrm{d}\eta. \tag{4.38}$$

The expression (2.6) shows that $J\mathbb{J}^{-1}$ is a matrix of first-order polynomials in (ξ, η) coordinates. Therefore, we have

$$\int_K |\nabla v|^2 \mathrm{d}x\mathrm{d}y = \int_\square J_K J^{-1} |\mathbb{J}_K^{-1} J \mathbb{J}^{-1}\widetilde{\nabla}\tilde{v}|^2 \mathrm{d}\xi\mathrm{d}\eta \le 4M \ln 2,$$

by using (4.38) and (2.8). Here, $M := \max_{\overline{\square}} 16 J_K |\mathbb{J}_K^{-1} J \mathbb{J}^{-1}\widetilde{\nabla}\tilde{v}|^2$ is the bound of a polynomial on $[-1, 1]^2$. This implies $V_h^p(K) \subset H^1(K)$.

From the chain rule, we have

$$\nabla v = \mathbb{J}_K^{-1}\widehat{\nabla}\hat{v} = \mathbb{J}_K^{-1}\mathbb{J}^{-1}\widetilde{\nabla}\tilde{v}.$$

As shown above, $J\mathbb{J}^{-1}$ is a matrix of linear polynomials in (ξ, η) coordinates. Thus, $J\mathbb{J}_K^{-1}\mathbb{J}^{-1}\widetilde{\nabla}\tilde{v} \in (\mathcal{Q}_p(\square))^2$ for any $v \in V_h^p(K)$, which implies $(J \circ T \circ \mathscr{F}_K)\nabla v \in \boldsymbol{V}_h^p(K)$. In the analysis of the HDG method, we know that ∇v belongs to the corresponding vector polynomial space as v is a polynomial in the physical domain. This motivates us to introduce weighted projections in the error analysis below.

According to (2.7), $T : \square \mapsto \triangle$ is a linear mapping when it is restricted to the elemental edges, so $Y_p(\triangle)|_{\partial\triangle}$ consists of piecewise pth-order polynomials. After the affine mapping $\mathscr{F}_K$, we have $V_h^p(K)|_{\partial K}$ is a piecewise pth-order polynomial space too. As all pulled edges are considered as two separate edges in $\widetilde{\mathcal{E}}_h$, piecewise

polynomial space $\mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$ confined on a pulled edge is still a piecewise pth-order polynomial space. Therefore, we always have $V_h^p(K)|_{\partial K} \subset \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)|_{\partial K}$. □

As $V_h^p(K)$ and $\boldsymbol{V_h^p}(K)$ consist of non-polynomial basis functions, we do not have $\nabla v \in \boldsymbol{V_h^p}(K)$ for all $v \in V_h^p(K)$. Instead, the conclusion (ii) in Proposition 1 holds. Therefore, the standard L^2-projection is not suitable for the error analysis. In light of this, we introduce the following J^{-1}-weighted projections. Let $\widetilde{\Pi}^p$ and $\widetilde{\boldsymbol{\Pi}}^p$ be the standard L^2-projections from $L^2(\square)$ to $\mathcal{Q}_p(\square)$ and $\boldsymbol{L}^2(\square) := (L^2(\square))^2$ to $\boldsymbol{Q}_p(\square) := (\mathcal{Q}_p(\square))^2$, respectively. We define the elemental operators:

$$\begin{aligned} &\Pi_K^p : L^2(K) \to V_h^p(K), \quad \Pi_K^p u = (\widetilde{\Pi}^p \tilde{u}) \circ T^{-1} \circ \mathscr{F}_K^{-1}, \\ &\boldsymbol{\Pi}_K^p : \boldsymbol{L}^2(K) \to \boldsymbol{V_h^p}(K), \quad \boldsymbol{\Pi}_K^p \boldsymbol{q} = (\widetilde{\boldsymbol{\Pi}}^p \tilde{\boldsymbol{q}}) \circ T^{-1} \circ \mathscr{F}_K^{-1}, \end{aligned} \tag{4.39}$$

for all $K \in \mathcal{T}_h$. They are actually weighted projection operators as one can verify that

$$(u - \Pi_K^p u, w)_{J^{-1},K} = |J_K| \int_{\square} (\tilde{u} - \widetilde{\Pi}^p \tilde{u}) \tilde{w} \mathrm{d}\xi \mathrm{d}\eta = 0, \quad \forall w \in V_h^p(K), \tag{4.40}$$

where the weight J^{-1} is the inverse of the Jacobian defined in (2.3). Accordingly, we define the piecewise projection operators:

$$\begin{aligned} &\Pi_h^p : L^2(\mathcal{T}_h) \to V_h^p(\mathcal{T}_h), \quad (\Pi_h^p u)|_K = \Pi_K^p (u|_K), \quad \forall K \in \mathcal{T}_h, \\ &\boldsymbol{\Pi}_h^p : \boldsymbol{L}^2(\mathcal{T}_h) \to \boldsymbol{V_h^p}(\mathcal{T}_h), \quad (\boldsymbol{\Pi}_h^p \boldsymbol{q})|_K = \boldsymbol{\Pi}_K^p (\boldsymbol{q}|_K), \quad \forall K \in \mathcal{T}_h. \end{aligned} \tag{4.41}$$

Then, the orthogonality of the standard L^2-projection operator in the reference (ξ, η) coordinates implies:

Proposition 2 *If $u \in L^2(\mathcal{T}_h)$ and $\boldsymbol{q} \in \boldsymbol{L}^2(\mathcal{T}_h)$, then*

$$\begin{aligned} &(u - \Pi_h^p u,\ w)_{\mathcal{T}_h} &&= 0, \quad \forall\ w \in V_h^{p-1}(\mathcal{T}_h); \\ &(u - \Pi_h^p u,\ \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} &&= 0, \quad \forall\ \boldsymbol{v} \in \boldsymbol{V_h^p}(\mathcal{T}_h); \\ &(\boldsymbol{q} - \boldsymbol{\Pi_h^p} \boldsymbol{q},\ \nabla w)_{\mathcal{T}_h} &&= 0, \quad \forall\ w \in V_h^p(\mathcal{T}_h). \end{aligned} \tag{4.42}$$

Proof We only need to verify the local projection of the first two equations and the third one can be obtained in the same fashion as for the second one. The expression (2.3) and the definition of $V_h^p(K)$ imply that $J\tilde{w} \in \mathcal{Q}_p(\square)$ for all $w \in V_h^{p-1}(K)$. Thus,

$$(u - \Pi_K^p u,\ w)_K = J_K(\tilde{u} - \widetilde{\Pi}^p \tilde{u},\ J\tilde{w})_{\square} = 0.$$

From the expression (2.6), we have

$$J\mathbb{J}^{-1} = \frac{1}{8} \begin{bmatrix} 3 - \xi & \eta + 1 \\ \xi + 1 & 3 - \eta \end{bmatrix}, \tag{4.43}$$

which implies $(J\mathbb{J}^{-1}\widetilde{\nabla}) \cdot \tilde{\boldsymbol{v}} \in \mathcal{Q}_p(\square)$ for all $\boldsymbol{v} \in \boldsymbol{V_h^p}(K)$. Note that $\nabla \cdot \boldsymbol{v} = (\mathbb{J}_K^{-1}\mathbb{J}^{-1}\widetilde{\nabla}) \cdot \tilde{\boldsymbol{v}}$, then

$$(u - \Pi_K^p u,\ \nabla \cdot \boldsymbol{v})_K = J_K\big(\tilde{u} - \widetilde{\Pi}^p \tilde{u},\ (\mathbb{J}_K^{-1} J \mathbb{J}^{-1} \widetilde{\nabla}) \cdot \tilde{\boldsymbol{v}}\big)_{\square} = 0. \tag{4.44}$$

This completes the proof. □

By the assumption that $\mathcal{T}_h$ is a shape regular *quasi-uniform* mesh of the domain Ω, we have the following two lemmas.

Lemma 1 (see [5]) *For any $K \in \mathcal{T}_h$ and $v \in H^s(K)$ with $s \geq 0$,*

$$|v|_{s,K} \leq Ch^{1-s}|\hat{v}|_{s,\triangle}, \quad |\hat{v}|_{s,\triangle} \leq Ch^{s-1}|v|_{s,K}, \tag{4.45}$$

where C is a generic constant independent of h.

Lemma 2 (see [13]) *For any $K \in \mathcal{T}_h$ and $u \in H^{s+1}(K)$ with $s \geq 0$, there exists a polynomial $u_p \in \mathcal{P}_p(K)$ such that*

$$\frac{p}{h}\|u-u_p\|_K + \sqrt{\frac{p}{h}}\|u-u_p\|_{\partial K} + \|\nabla(u-u_p)\|_K \leq C_s\, h^{\min\{p,s\}} p^{-s}|u|_{s+1,K},$$

where C_s is a generic constant depending only on s.

The estimates for the weighted projections Π_h^p and $\mathbf{\Pi}_h^p$ are the key ingredients for the error analysis. As they are piecewisely defined on each element K, we conduct the error analysis on each element.

Lemma 3 *For any $K \in \mathcal{T}_h$, if $u \in H^1(K)$, then*

$$\|u - \Pi_K^p u\|_K \leq Ch\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_\square \leq Chp^{-1}|u|_{1,K}. \tag{4.46}$$

Proof Using the transform from the triangle K to the reference square $\square$ gives

$$\|u - \Pi_K^p u\|_K = \sqrt{J_K}\|\sqrt{J}(\tilde{u} - \widetilde{\Pi}^p\tilde{u})\|_\square \leq Ch\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_\square. \tag{4.47}$$

For the error estimate of the standard L^2-projection, we recall the result directly obtained from [18, Lemma 3.4] and the Stirling's formula, i.e.,

$$\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_\square \leq Cp^{-1}\big(\|(1-\xi^2)^{\frac{1}{2}}\partial_\xi\tilde{u}\|_\square + \|(1-\eta^2)^{\frac{1}{2}}\partial_\eta\tilde{u}\|_\square\big). \tag{4.48}$$

Noting that

$$\begin{aligned}
\partial_\xi\tilde{u} &= \frac{\partial\hat{x}}{\partial\xi}\partial_{\hat{x}}\hat{u} + \frac{\partial\hat{y}}{\partial\xi}\partial_{\hat{y}}\hat{u} = \frac{3-\eta}{8}\partial_{\hat{x}}\hat{u} - \frac{1+\eta}{8}\partial_{\hat{y}}\hat{u},\\
\partial_\eta\tilde{u} &= \frac{\partial\hat{x}}{\partial\eta}\partial_{\hat{x}}\hat{u} + \frac{\partial\hat{y}}{\partial\eta}\partial_{\hat{y}}\hat{u} = \frac{3-\xi}{8}\partial_{\hat{y}}\hat{u} - \frac{1+\xi}{8}\partial_{\hat{x}}\hat{u},
\end{aligned} \tag{4.49}$$

then

$$\begin{aligned}
\|(1-\xi^2)^{\frac{1}{2}}\partial_\xi\tilde{u}\|_\square^2 &= \int_\square (1-\xi^2)|\partial_\xi\tilde{u}|^2 \mathrm{d}\xi\mathrm{d}\eta\\
&= \int_\square (1-\xi^2)\left|\frac{3-\eta}{8}\partial_{\hat{x}}\hat{u} - \frac{1+\eta}{8}\partial_{\hat{y}}\hat{u}\right|^2 \mathrm{d}\xi\mathrm{d}\eta\\
&= \int_\triangle \frac{(1-\xi^2)}{J}\left|\frac{3-\eta}{8}\partial_{\hat{x}}\hat{u} - \frac{1+\eta}{8}\partial_{\hat{y}}\hat{u}\right|^2 \mathrm{d}\hat{x}\mathrm{d}\hat{y}.
\end{aligned} \tag{4.50}$$

For all $(\xi,\eta)\in\square$, we have $1-\xi\le 2-\xi-\eta=16J$. Therefore, $(1-\xi^2)/J$, $(3-\eta)/8$ and $(1+\eta)/8$ are bounded for all $(\xi,\eta)\in\square$. Together with Lemma 1, we arrive at

$$\|(1-\xi^2)^{\frac{1}{2}}\partial_\xi\tilde{u}\|_\square^2\le C|\hat{u}|_{1,\triangle}^2\le C|u|_{1,K}^2. \tag{4.51}$$

By swapping $\hat{x}\leftrightarrow\hat{y}$ and $\xi\leftrightarrow\eta$, we also have

$$\|(1-\eta^2)^{\frac{1}{2}}\partial_\eta\tilde{u}\|_\square^2\le C|u|_{1,K}^2. \tag{4.52}$$

A combination of (4.47)–(4.48) and (4.51)–(4.52) leads to the desired result. □

The following two lemmas will be useful in deriving the error bound for the projection $\Pi_K^p u$ on the boundary of K.

Lemma 4 *Given $u\in H^1(K)$ and vector $\boldsymbol{\xi}=(\xi,\eta)$, then*

$$\|\widetilde{\nabla}\tilde{u}\cdot\boldsymbol{\xi}\|_\square\le C|u|_{1,K}. \tag{4.53}$$

Proof Recalling that

$$\frac{1}{J(\xi,\eta)}=\frac{16}{2-\xi-\eta}\ge 4,\quad\forall(\xi,\eta)\in\square,$$

we have

$$\|\widetilde{\nabla}\tilde{u}\cdot\boldsymbol{\xi}\|_\square^2\le\|\widetilde{\nabla}\tilde{u}\cdot\boldsymbol{\xi}\|_{J^{-1},\square}^2=\int_\square\left|\xi\partial_\xi\tilde{u}+\eta\partial_\eta\tilde{u}\right|^2J^{-1}\mathrm{d}\xi\mathrm{d}\eta. \tag{4.54}$$

On the other hand, one can verify that the matrix

$$\begin{aligned}\mathbb{B}&=\begin{bmatrix}3-\xi & 1+\eta\\ 1+\xi & 3-\eta\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}3-\xi & 1+\eta\\ 1+\xi & 3-\eta\end{bmatrix}-\begin{bmatrix}\xi^2 & \xi\eta\\ \xi\eta & \eta^2\end{bmatrix}\\&=\begin{bmatrix}6+(2-\xi)^2 & 6+2\xi+2\eta-3\xi\eta\\ 6+2\xi+2\eta-3\xi\eta & 6+(2-\eta)^2\end{bmatrix},\end{aligned}$$

is a positive semidefinite matrix for all $(\xi,\eta)\in\square$. In fact, its diagonal entries are positive and its determinant

$$\det(\mathbb{B})=6(\xi-\eta)^2+4(1-\xi)(1-\eta)(16-2\xi\eta), \tag{4.55}$$

which is non-negative for all $(\xi,\eta)\in\square$. Therefore, it holds $(\widetilde{\nabla}\tilde{u})^{\mathrm{T}}\mathbb{B}(\widetilde{\nabla}\tilde{u})\ge 0$, i.e,

$$(\widetilde{\nabla}\tilde{u})^{\mathrm{T}}\begin{bmatrix}3-\xi & 1+\eta\\ 1+\xi & 3-\eta\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}3-\xi & 1+\eta\\ 1+\xi & 3-\eta\end{bmatrix}\widetilde{\nabla}\tilde{u}-(\widetilde{\nabla}\tilde{u})^{\mathrm{T}}\begin{bmatrix}\xi\\ \eta\end{bmatrix}\begin{bmatrix}\xi & \eta\end{bmatrix}\widetilde{\nabla}\tilde{u}\ge 0, \tag{4.56}$$

for all $(\xi,\eta)\in\square$, or equivalently

$$\left|\xi\partial_\xi\tilde{u}+\eta\partial_\eta\tilde{u}\right|^2\le\left|(3-\xi)\partial_\xi\tilde{u}+(1+\eta)\partial_\eta\tilde{u}\right|^2+\left|(1+\xi)\partial_\xi\tilde{u}+(3-\eta)\partial_\eta\tilde{u}\right|^2.$$

Together with the expression

$$\widehat{\nabla}\hat{u}=\mathbb{J}^{-1}\widetilde{\nabla}\tilde{u}=\frac{1}{8J}\begin{bmatrix}(3-\xi)\partial_\xi\tilde{u}+(1+\eta)\partial_\eta\tilde{u}\\ (1+\xi)\partial_\xi\tilde{u}+(3-\eta)\partial_\eta\tilde{u}\end{bmatrix}, \tag{4.57}$$

we obtain

$$\left|\xi\partial_\xi\tilde{u}+\eta\partial_\eta\tilde{u}\right|^2J^{-1}\le 64J|\widehat{\nabla}\hat{u}|^2,\quad\forall(\xi,\eta)\in\square. \tag{4.58}$$

Substituting the above estimate into (4.54) and applying Lemma 1 leads to

$$\|\widetilde{\nabla}\tilde{u}\cdot\boldsymbol{\xi}\|_{\square}^2 \le 64|\hat{u}|_{1,\triangle}^2 \le C|u|_{1,K}^2. \tag{4.59}$$

This completes the proof. □

Lemma 5 *If $v \in H^1(K) \cap L^2_{J^{-1}}(K)$, then*

$$\|\tilde{v}\|_{\partial\square}^2 = 2\|\tilde{v}\|_{\square}^2 + 2(\tilde{v}, \widetilde{\nabla}\tilde{v}\cdot\boldsymbol{\xi})_{\square}, \tag{4.60}$$

where $\boldsymbol{\xi} = (\xi, \eta)$.

Proof The result can be obtained by mimic the proof of Lemma 2.19 in [10]. The only change here is that we have

$$\boldsymbol{\xi}\cdot\boldsymbol{n} = 1 \quad \text{on} \quad \partial\square,$$

where $\boldsymbol{n}$ is the unit outer normal of $\partial\square$. The identity (4.60) follows from

$$\|\tilde{v}\|_{\partial\square}^2 = \int_{\partial\square}\left(\tilde{v}^2\boldsymbol{\xi}\right)\cdot\boldsymbol{n}\mathrm{d}S = \int_{\square}\widetilde{\nabla}\cdot\left(\tilde{v}^2\boldsymbol{\xi}\right)\mathrm{d}\xi\mathrm{d}\eta$$

and

$$\widetilde{\nabla}\cdot\left(\tilde{v}^2\boldsymbol{\xi}\right) = \tilde{v}^2\widetilde{\nabla}\cdot\boldsymbol{\xi} + 2\tilde{v}\widetilde{\nabla}\tilde{v}\cdot\boldsymbol{\xi} = 2\tilde{v}^2 + 2\tilde{v}\widetilde{\nabla}\tilde{v}\cdot\boldsymbol{\xi}.$$

The assumption $v \in H^1(K) \cap L^2_{J^{-1}}(K)$ ensures that the integral

$$\begin{aligned}\left|\int_{\square}\widetilde{\nabla}\cdot\left(\tilde{v}^2\boldsymbol{\xi}\right)\mathrm{d}\xi\mathrm{d}\eta\right| &\le 2\|\tilde{v}\|_{\square}(\|\tilde{v}\|_{\square}+1) + 2\|\widetilde{\nabla}\tilde{v}\cdot\boldsymbol{\xi}\|_{\square}\\ &\le C(\|v\|_{J^{-1},K}^2 + \|v\|_{J^{-1},K} + |v|_{1,K}),\end{aligned} \tag{4.61}$$

is finite. □

Remark 1 Noting that $J^{-1}(\xi,\eta)$ is singular only at a boundary point and the integral (2.8) is finite, the assumptions $L^2_{J^{-1}}(K)$ can be satisfied if $u \in L^\infty(K)$.

Lemma 6 *If $u \in H^1(K) \cap L^2_{J^{-1}}(K)$, then*

$$\|u - \Pi_K^p u\|_{\partial K} \le Ch^{\frac{1}{2}}p^{-\frac{1}{2}}|u|_{1,K}.$$

Proof As mentioned in (2.7), $T : \square \to \triangle$ is a piecewise affine mapping when it is confined to the boundary $\partial\square$. Therefore,

$$\|u - \Pi_K^p u\|_{\partial K}^2 \le Ch\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_{\partial\square}^2. \tag{4.62}$$

By Lemma 5, the boundary error can be written as

$$\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_{\partial\square}^2 = 2\|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_{\square}^2 + 2(\tilde{u} - \widetilde{\Pi}^p\tilde{u}, \widetilde{\nabla}(\tilde{u} - \widetilde{\Pi}^p\tilde{u})\cdot\boldsymbol{\xi})_{\square}. \tag{4.63}$$

Noting that $\widetilde{\nabla}(\widetilde{\Pi}^p\tilde{u})\cdot\boldsymbol{\xi} \in \mathcal{Q}_p(\square)$, we obtain

$$\begin{aligned}(\tilde{u} - \widetilde{\Pi}^p\tilde{u}, \widetilde{\nabla}(\tilde{u} - \widetilde{\Pi}^p\tilde{u})\cdot\boldsymbol{\xi})_{\square} &= (\tilde{u} - \widetilde{\Pi}^p\tilde{u}, \nabla\tilde{u}\cdot\boldsymbol{\xi})_{\square}\\ &\le \|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_{\square}\|\nabla\tilde{u}\cdot\boldsymbol{\xi}\|_{\square} \le \|\tilde{u} - \widetilde{\Pi}^p\tilde{u}\|_{\square}|u|_{1,K},\end{aligned} \tag{4.64}$$

by using the orthogonality of $\widetilde{\Pi}^p$, the Cauchy-Schwarz's inequality and Lemma 4. Therefore, (4.62)–(4.64) gives the estimate

$$\|u - \Pi_K^p u\|_{\partial K}^2 \le Ch(\|\tilde{u} - \widetilde{\Pi}^p \tilde{u}\|_{\square}^2 + \|\tilde{u} - \widetilde{\Pi}^p \tilde{u}\|_{\square} |u|_{1,K}).$$

Using the estimate $\|\tilde{u} - \widetilde{\Pi}^p \tilde{u}\|_{\square} \le Cp^{-1}|u|_{1,K}$ in Lemma 3 completes the proof. □

With the results in Lemmas 3 and 6, we can summarize the error estimates for the projection operators Π_h^p and $\mathbf{\Pi}_h^p$ as follows.

Theorem 1 *Suppose* $u|_K \in H^{p+1}(K) \cap L^2_{J^{-1}}(K)$ *and* $\boldsymbol{q}|_K \in (H^p(K))^2 \cap \left(L^2_{J^{-1}}(K)\right)^2$ *hold for any* $K \in \mathcal{T}_h$. *Then*

$$\|u - \Pi_h^p u\|_{\mathcal{T}_h} \le Ch^s p^{-s} |u|_{s,\mathcal{T}_h}, \qquad 1 \le s \le p+1, \tag{4.65a}$$

$$\|\boldsymbol{q} - \mathbf{\Pi}_h^p \boldsymbol{q}\|_{\mathcal{T}_h} \le Ch^r p^{-r} |\boldsymbol{q}|_{r,\mathcal{T}_h}, \qquad 1 \le r \le p, \tag{4.65b}$$

$$\|u - \Pi_h^p u\|_{\partial\mathcal{T}_h} \le Ch^{s-\frac{1}{2}} p^{\frac{1}{2}-s} |u|_{s,\mathcal{T}_h}, \qquad 1 \le s \le p+1, \tag{4.65c}$$

$$\|\boldsymbol{q} - \mathbf{\Pi}_h^p \boldsymbol{q}\|_{\partial\mathcal{T}_h} \le Ch^{r-\frac{1}{2}} p^{\frac{1}{2}-r} |\boldsymbol{q}|_{r,\mathcal{T}_h}, \qquad 1 \le r \le p. \tag{4.65d}$$

Proof We only consider the error estimates (4.65a) and (4.65c). The other two can be obtained similarly. By Lemma 2 and Proposition 1, there exists $u_p \in \mathcal{P}_p(K) \subset V_h^p(K)$ such that

$$|u - u_p|_{1,K} \le Ch^{s-1} p^{1-s} |u|_{s,K}, \quad \forall\, 1 \le s \le p+1. \tag{4.66}$$

Apparently, $\tilde{u}_p \in H^1(\square)$ as $\tilde{u}_p$ is still a polynomial on $\square$. By using the fact $\Pi_K^p w = w$ for all $w \in V_h^p(K)$ and the estimates in Lemma 3 and (4.66), we obtain

$$\begin{aligned} \|u - \Pi_K^p u\|_K &= \|(u - u_p) - \Pi_K^p (u - u_p)\|_K \\ &\le Chp^{-1} |u - u_p|_{1,K} \le Ch^s p^{-s} |u|_{s,K}, \end{aligned} \tag{4.67}$$

for all $1 \le s \le p+1$.

Using Lemma 6 and following the above proof, we can obtain

$$\begin{aligned} \|u - \Pi_K^p u\|_{\partial K} &= \|(u - u_p) - \Pi_K^p (u - u_p)\|_{\partial K} \\ &\le Ch^{\frac{1}{2}} p^{-\frac{1}{2}} |u - u_p|_{1,K} \le Ch^{s-\frac{1}{2}} p^{\frac{1}{2}-s} |u|_{s,K}, \end{aligned} \tag{4.68}$$

for all $1 \le s \le p+1$. By the quasi-uniform mesh assumption, we can collect the estimates (4.67)–(4.68) to obtain (4.65a) and (4.65c). □

Denote by $\mathcal{P}_p(e)$ the set of all polynomials of degree at most p on any edge $e \in \widetilde{\mathcal{E}}_h$. Let $\mathbf{P}_e^p$ be the L^2 projection from $L^2(e)$ to $\mathcal{P}_p(e)$, for any $e \in \widetilde{\mathcal{E}}_h$. Define the global projection operator $\mathbf{P}_\partial^p$ as follows

$$\mathbf{P}_\partial^p : L^2(\widetilde{\mathcal{E}}_h) \to \mathcal{M}_h^P(\widetilde{\mathcal{E}}_h) \quad \text{such that} \quad (\mathbf{P}_\partial^p u)|_e = \mathbf{P}_e^p(u|_e), \quad \forall e \in \widetilde{\mathcal{E}}_h.$$

By the conclusion (iii) in Proposition 1, we have the following orthogonality,

Lemma 7 *Given $u \in L^2(\widetilde{\mathcal{E}}_h)$, we have*

$$\begin{aligned} \langle u - \mathbf{P}_\partial^p u,\ \mu_h^p \rangle_{\widetilde{\mathcal{E}}_h} = 0, \quad \forall \mu_h^p \in \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h), \\ \langle u - \mathbf{P}_\partial^p u,\ v_h^p \rangle_{\partial \mathcal{T}_h} = 0, \quad \forall v_h^p \in V_h^p(\mathcal{T}_h). \end{aligned} \tag{4.69}$$

The standard L^2-error estimate for this projection is stated as follows.

Lemma 8 *Suppose $u|_K \in H^{s+1}(K)$ and $\boldsymbol{q}|_K \in (H^s(K))^2$ hold for any $K \in \mathcal{T}_h$. Then*

$$\|u - \mathbf{P}_\partial^p u\|_{\partial \mathcal{T}_h} \leq C h^{s-\frac{1}{2}} p^{\frac{1}{2}-s} |u|_{s,\mathcal{T}_h}, \quad 1 \leq s \leq p+1, \tag{4.70a}$$

$$\|\boldsymbol{q} \cdot \boldsymbol{n} - \mathbf{P}_\partial^p \boldsymbol{q} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} \leq C h^{r-\frac{1}{2}} p^{\frac{1}{2}-r} |\boldsymbol{q}|_{r,\mathcal{T}_h}, \quad 1 \leq r \leq p+1. \tag{4.70b}$$

In the rest of this section, we will take three steps to prove the L^2-error estimate of the HDTSEM solutions. For simplicity, we take $\beta(x) \equiv 1$ and $\gamma(x) \equiv 0$ in the analysis. The main theoretical results are presented in Theorems 3 and 4.

Step 1: The error equation Define

$$\begin{aligned} &e_u := \Pi_h^p u - u_h^p, \quad e_{\boldsymbol{q}} := \boldsymbol{\Pi}_h^p \boldsymbol{q} - \boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}, \quad \hat{e}_u := \mathbf{P}_\partial^p u - \hat{u}_h^p, \\ &\delta_u := u - \Pi_h^p u, \quad \delta_{\boldsymbol{q}} := \boldsymbol{q} - \boldsymbol{\Pi}_h^p \boldsymbol{q}. \end{aligned}$$

Lemma 9 *Let $(u, \boldsymbol{q})$, $(u_h^p, \boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}, \hat{u}_h^p)$ solve (3.21) and (3.24a), respectively, with the coefficients $\beta(x) \equiv 1$, $\gamma(x) \equiv 0$. Then,*

$$(e_{\boldsymbol{q}}, \boldsymbol{v})_{\mathcal{T}_h} - (e_u, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} + \langle \hat{e}_u, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = -\left(\delta_{\boldsymbol{q}}, \boldsymbol{v}\right)_{\mathcal{T}_h}, \tag{4.71a}$$

$$-(e_{\boldsymbol{q}}, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q} \cdot \boldsymbol{n} - \hat{\boldsymbol{q}}_h^p \cdot \boldsymbol{n}, w \rangle_{\partial \mathcal{T}_h} = 0, \tag{4.71b}$$

$$\langle \boldsymbol{q} \cdot \boldsymbol{n} - \hat{\boldsymbol{q}}_h^p \cdot \boldsymbol{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0, \tag{4.71c}$$

$$\langle \hat{e}_u, \mu \rangle_{\partial \Omega} = 0, \tag{4.71d}$$

for all $(w, \boldsymbol{v}, \mu) \in V_h^p(\mathcal{T}_h) \times \boldsymbol{V}_{\boldsymbol{h}}^{\boldsymbol{p}}(\mathcal{T}_h) \times \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$. In addition, on $\partial \mathcal{T}_h$, we have the following characterization:

$$\boldsymbol{q} \cdot \boldsymbol{n} - \hat{\boldsymbol{q}}_h^p \cdot \boldsymbol{n} = e_{\boldsymbol{q}} \cdot \boldsymbol{n} + \tau \frac{p}{h} (e_u - \hat{e}_u) + \delta_{\boldsymbol{q}} \cdot \boldsymbol{n} + \tau \frac{p}{h} \mathbf{P}_\partial^p \delta_u. \tag{4.72}$$

Proof Firstly, we note that the exact solution $(u, \boldsymbol{q}, u|_{\partial \mathcal{T}_h})$ and the trace $\boldsymbol{q} \cdot \boldsymbol{n}$ satisfy the equations in (3.24a). Hence, by the orthogonal properties (4.42) and (4.69), we can replace the unknowns by their projections and do some simple algebraic manipulations to obtain

$$(\boldsymbol{\Pi}_h^p \boldsymbol{q}, \boldsymbol{v})_{\mathcal{T}_h} - (\Pi_h^p u, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} + \langle \mathbf{P}_\partial^p u, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = \left(\boldsymbol{\Pi}_{\boldsymbol{h}}^{\boldsymbol{p}} \boldsymbol{q} - \boldsymbol{q}, \boldsymbol{v}\right)_{\mathcal{T}_h}, \tag{4.73a}$$

$$-(\boldsymbol{\Pi}_h^p \boldsymbol{q}, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q} \cdot \boldsymbol{n}, w \rangle_{\partial \mathcal{T}_h} = (f, w)_{\mathcal{T}_h}, \tag{4.73b}$$

$$\langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0, \tag{4.73c}$$

$$\langle \mathbf{P}_\partial^p u, \mu \rangle_{\partial \Omega} = 0, \tag{4.73d}$$

for all $(w, \boldsymbol{v}, \mu) \in V_h^p(\mathcal{T}_h) \times \boldsymbol{V}_{\boldsymbol{h}}^{\boldsymbol{p}}(\mathcal{T}_h) \times \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)$. The error equations (4.71a) are obtained by subtracting the equations (3.24a) from (4.73a), respectively.

Recalling the conclusion $V_h^p(K)|_{\partial K} \subset \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)|_{\partial K}$ in Proposition 1, we have $\Pi_h^p u|_{\partial K} \in \mathcal{M}_h^p(\widetilde{\mathcal{E}}_h)|_{\partial K}$. Hence, $\mathbf{P}_\partial^p(\Pi_h^p u|_{\partial\mathcal{T}_h}) = \Pi_h^p u|_{\partial\mathcal{T}_h}$ and

$$\hat{u}_h^p - u_h^p = \hat{u}_h^p - \mathbf{P}_\partial^p u + \mathbf{P}_\partial^p u - \mathbf{P}_\partial^p \Pi_h^p u + \mathbf{P}_\partial^p \Pi_h^p u - u_h^p = -\hat{e}_u + \mathbf{P}_\partial^p \delta_u + e_u \quad \text{on } \partial\mathcal{T}_h.$$

Together with the expression (3.25), we get

$$\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p \cdot \boldsymbol{n} = \boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{q}_h^p\cdot\boldsymbol{n} - \tau\frac{p}{h}\left(u_h^p - \hat{u}_h^p\right) = e_{\boldsymbol{q}}\cdot\boldsymbol{n} + \delta_{\boldsymbol{q}}\cdot\boldsymbol{n} + \tau\frac{p}{h}\left(e_u - \hat{e}_u\right) + \tau\frac{p}{h}\mathbf{P}_\partial^p\delta_u.$$

This ends the proof. □

Step 2: Estimate of $\boldsymbol{e_q}$ To present error estimate for $e_{\boldsymbol{q}}$, we first define the energy norm of the error as

$$|||e_u, e_{\boldsymbol{q}}, \hat{e}_u|||^2 := \|e_{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \tau h^{-\frac{1}{2}} p^{\frac{1}{2}} \|e_u - \hat{e}_u\|_{\partial\mathcal{T}_h}^2. \tag{4.74}$$

Theorem 2 *Suppose that the exact solution of (3.21) satisfies $u|_K \in H^{s+1}(K) \cap L^2_{J^{-1}}(K)$ and $\boldsymbol{q}|_K \in (H^s(K))^2 \cap \left(L^2_{J^{-1}}(K)\right)^2$ for any $K \in \mathcal{T}_h$. Then*

$$\begin{aligned} |||e_u, e_{\boldsymbol{q}}, \hat{e}_u|||^2 &\le Ch^{\frac{1}{2}} p^{-\frac{1}{2}} \|\delta_{\boldsymbol{q}}\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_h} + Ch^{-\frac{1}{2}} p^{\frac{1}{2}} \|\delta_u\|_{\partial\mathcal{T}_h} + \|\delta_{\boldsymbol{q}}\|_{\mathcal{T}_h} \\ &\le Ch^s p^{-s}(|u|_{s+1,\mathcal{T}_h} + |\boldsymbol{q}|_{s,\mathcal{T}_h}), \end{aligned} \tag{4.75}$$

for all $1 \le s \le p$.

Proof Taking $(w, \boldsymbol{v}, \mu) = (e_u, e_{\boldsymbol{q}}, \hat{e}_u)$ in the error equations (4.71a) gives

$$(e_{\boldsymbol{q}}, e_{\boldsymbol{q}})_{\mathcal{T}_h} - (e_u, \nabla\cdot e_{\boldsymbol{q}})_{\mathcal{T}_h} + \left\langle \hat{e}_u, e_{\boldsymbol{q}}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} = -\left(\delta_{\boldsymbol{q}}, e_{\boldsymbol{q}}\right)_{\mathcal{T}_h}, \tag{4.76a}$$

$$-(e_{\boldsymbol{q}}, \nabla e_u)_{\mathcal{T}_h} + \left\langle \boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, e_u\right\rangle_{\partial\mathcal{T}_h} = 0, \tag{4.76b}$$

$$\left\langle \boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \hat{e}_u\right\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{4.76c}$$

$$\left\langle \hat{e}_u, \hat{e}_u\right\rangle_{\partial\Omega} = 0. \tag{4.76d}$$

Adding up the equations (4.76a) and (4.76b) and applying integration by parts, we get

$$(e_{\boldsymbol{q}}, e_{\boldsymbol{q}})_{\mathcal{T}_h} + \left\langle \hat{e}_u - e_u, e_{\boldsymbol{q}}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} + \left\langle \boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, e_u\right\rangle_{\partial\mathcal{T}_h} = -\left(\delta_{\boldsymbol{q}}, e_{\boldsymbol{q}}\right)_{\mathcal{T}_h}.$$

Moreover, the equations (4.76c) and (4.76d) imply

$$\left\langle \boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \hat{e}_u\right\rangle_{\partial\mathcal{T}_h} = 0.$$

Thus,

$$\|e_{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \left\langle \boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n} - e_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad e_u - \hat{e}_u\right\rangle_{\partial\mathcal{T}_h} = -(\delta_{\boldsymbol{q}}, \quad e_{\boldsymbol{q}})_{\mathcal{T}_h}. \tag{4.77}$$

Substituting the characterization (4.72) into (4.77) and rearranging resulted terms gives

$$|e_u, e_{\boldsymbol{q}}, \hat{e}_u|^2 = -\left\langle \delta_{\boldsymbol{q}} \cdot n + \tau \frac{p}{h} \mathbf{P}_\partial^p \delta_u,\ e_u - \hat{e}_u \right\rangle_{\partial \mathcal{T}_h} - (\delta_{\boldsymbol{q}},\ e_{\boldsymbol{q}})_{\mathcal{T}_h}. \tag{4.78}$$

Next, we will give estimate for the right-hand side term-by-term. By the definition (4.74) and Cauchy-Schwarz inequality, we have

$$\begin{aligned} \left| \left\langle \delta_{\boldsymbol{q}} \cdot \boldsymbol{n},\quad e_u - \hat{e}_u \right\rangle_{\partial \mathcal{T}_h} \right| &\le \|\delta_{\boldsymbol{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} \|e_u - \hat{e}_u\|_{\partial \mathcal{T}_h} \\ &\le C h^{\frac{1}{2}} p^{-\frac{1}{2}} \|\delta_{\boldsymbol{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} |\!|\!| e_u, e_{\boldsymbol{q}}, \hat{e}_u |\!|\!|. \end{aligned} \tag{4.79}$$

Similarly,

$$\left\langle \tau \frac{p}{h} \mathbf{P}_\partial^p \delta_u,\quad e_u - \hat{e}_u \right\rangle_{\partial \mathcal{T}_h} \le C h^{-\frac{1}{2}} p^{\frac{1}{2}} \|\delta_u\|_{\partial \mathcal{T}_h} |\!|\!| e_u, e_{\boldsymbol{q}}, \hat{e}_u |\!|\!|, \tag{4.80}$$

and

$$(\delta_{\boldsymbol{q}},\quad e_{\boldsymbol{q}})_{\mathcal{T}_h} \le \|\delta_{\boldsymbol{q}}\|_{\mathcal{T}_h} \|e_{\boldsymbol{q}}\|_{\mathcal{T}_h} \le \|\delta_{\boldsymbol{q}}\|_{\mathcal{T}_h} |\!|\!| e_u, e_{\boldsymbol{q}}, \hat{e}_u |\!|\!|. \tag{4.81}$$

We complete the proof by using (4.79)–(4.81) in (4.78) and then applying the error estimates in Theorem 1. □

As a consequence of the above results, we obtain the h-p error estimate for $\boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}$ as follows.

Theorem 3 *Suppose that the exact solution of (3.21) satisfies* $u|_K \in H^{s+1}(K) \cap L^2_{J-1}(K)$, $\boldsymbol{q}|_K \in (H^s(K))^2 \cap \left(L^2_{J-1}(K)\right)^2$ *for any* $K \in \mathcal{T}_h$. *Then*

$$\|\boldsymbol{q} - \boldsymbol{q}_{\boldsymbol{h}}^{\boldsymbol{p}}\|_{\mathcal{T}_h} \le C h^s p^{-s} (|u|_{s+1,\mathcal{T}_h} + |\boldsymbol{q}|_{s,\mathcal{T}_h}), \quad 1 \le s \le p.$$

Step 3: Estimate for e_u We will adopt the duality argument to obtain the optimal error estimate for e_u. For this purpose, let $(\phi, \boldsymbol{\theta}) \in H^2(\Omega) \times H(\mathrm{div}, \Omega)$ be the solution of the adjoint problem:

$$\begin{aligned} \nabla \phi - \boldsymbol{\theta} &= 0 \quad \text{in } \Omega, \\ \nabla \cdot \boldsymbol{\theta} &= \eta \quad \text{in } \Omega, \\ \phi &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.82}$$

We assume the solution $(\phi, \boldsymbol{\theta})$ has the following elliptic regularity property:

$$\|\boldsymbol{\theta}\|_{1,\Omega} + \|\phi\|_{2,\Omega} \le C \|\eta\|_\Omega. \tag{4.83}$$

As the Sobolev embedding theorem gives $\|\phi\|_{\infty,\Omega} \le C\|\phi\|_{2,\Omega}$, Remark 1 implies that $\phi|_K \in L^2_{J-1}(K)$ for any $K \in \mathcal{T}_h$. In the following analysis, we further assume that $\boldsymbol{\theta}|_K \in (L^2_{J-1}(K))^2$ for any $K \in \mathcal{T}_h$. This assumption is acceptable as we will take $\eta = e_u$ which are piecewise polynomial on the mesh $\mathcal{T}_h$.

Lemma 10 *There holds*

$$\begin{aligned} \|e_u\|^2_{\mathcal{T}_h} &= (e_{\boldsymbol{q}} + \delta_{\boldsymbol{q}},\ \boldsymbol{\Pi}_h^p \boldsymbol{\theta} - \boldsymbol{\theta})_{\mathcal{T}_h} + \left\langle e_u - \hat{e}_u,\quad (\boldsymbol{\theta} - \boldsymbol{\Pi}_h^p \boldsymbol{\theta}) \cdot \boldsymbol{n} \right\rangle_{\partial \mathcal{T}_h} \\ &\quad - \langle \delta_{\boldsymbol{q}} \cdot \boldsymbol{n},\ \mathbf{P}_\partial^p \phi - \Pi_h^p \phi \rangle_{\partial \mathcal{T}_h} - \left\langle \tau \tfrac{p}{h} (e_u - \hat{e}_u),\ \ \phi - \Pi_h^p \phi \right\rangle_{\partial \mathcal{T}_h} \\ &\quad - \left\langle \tau \tfrac{p}{h} \mathbf{P}_\partial^p \delta_u,\ \ \phi - \Pi_h^p \phi \right\rangle_{\partial \mathcal{T}_h} + (\delta_{\boldsymbol{q}}, \nabla \phi)_{\mathcal{T}_h}. \end{aligned}$$

Proof Taking $\eta = e_u$ in the adjoint problem (4.82), we have

$$\|e_u\|^2_{\mathcal{T}_h} = (e_u, \quad \nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h} + (e_{\boldsymbol{q}}, \quad \nabla\phi)_{\mathcal{T}_h} - (e_{\boldsymbol{q}}, \quad \boldsymbol{\theta})_{\mathcal{T}_h}.$$

Inserting the projections $\Pi_h^p\phi$ and $\boldsymbol{\Pi}_h^p\boldsymbol{\theta}$, then applying integration by parts and the orthogonal properties in Lemma Proposition 2, we obtain

$$\begin{aligned}\|e_u\|^2_{\mathcal{T}_h} = \ & (e_u, \quad \nabla\cdot\boldsymbol{\Pi}_h^p\boldsymbol{\theta})_{\mathcal{T}_h} + (e_{\boldsymbol{q}}, \quad \nabla\Pi_h^p\phi)_{\mathcal{T}_h} - (e_{\boldsymbol{q}}, \quad \boldsymbol{\theta})_{\mathcal{T}_h} \\ & + \langle e_u, \quad (\boldsymbol{\theta} - \boldsymbol{\Pi}_h^p\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle e_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad \phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned} \tag{4.84}$$

On the other hand, taking $(w, \boldsymbol{v}) = (\Pi_h^p\phi, \boldsymbol{\Pi}_h^p\boldsymbol{\theta})$ in (4.71a), we have

$$\begin{aligned}(e_{\boldsymbol{q}} + \delta_{\boldsymbol{q}}, \boldsymbol{\Pi}_h^p\boldsymbol{\theta})_{\mathcal{T}_h} - (e_u, \nabla\cdot\boldsymbol{\Pi}_h^p\boldsymbol{\theta})_{\mathcal{T}_h} + \langle\hat{e}_u, \boldsymbol{\Pi}_h^p\boldsymbol{\theta}\cdot n\rangle_{\partial\mathcal{T}_h} &= 0, \\ -(e_{\boldsymbol{q}}, \nabla\Pi_h^p\phi)_{\mathcal{T}_h} + \langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} &= 0,\end{aligned} \tag{4.85}$$

which implies

$$\begin{aligned}(e_u, \nabla\cdot\boldsymbol{\Pi}_h^p\boldsymbol{\theta})_{\mathcal{T}_h} + (e_{\boldsymbol{q}}, \nabla\Pi_h^p\phi)_{\mathcal{T}_h} = \ & (e_{\boldsymbol{q}} + \delta_{\boldsymbol{q}}, \boldsymbol{\Pi}_h^p\boldsymbol{\theta})_{\mathcal{T}_h} + \langle\hat{e}_u, \boldsymbol{\Pi}_h^p\boldsymbol{\theta}\cdot n\rangle_{\partial\mathcal{T}_h} \\ & + \langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Substituting the above equation into (4.84) gives

$$\begin{aligned}\|e_u\|^2_{\mathcal{T}_h} = \ & (e_{\boldsymbol{q}} + \delta_{\boldsymbol{q}}, \quad \boldsymbol{\Pi}_h^p\boldsymbol{\theta} - \boldsymbol{\theta})_{\mathcal{T}_h} + \langle\hat{e}_u, \boldsymbol{\Pi}_h^p\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ & + \langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} + \langle e_u, (\boldsymbol{\theta} - \boldsymbol{\Pi}_h^p\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ & + \langle e_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad \phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} + (\delta_{\boldsymbol{q}}, \boldsymbol{\theta})_{\mathcal{T}_h}.\end{aligned} \tag{4.86}$$

From error (4.71d), we have $\hat{e}_u|_{\partial\Omega} = 0$. Together with the fact that $\hat{e}_u$ is single-valued across interior edges, we conclude that $\langle\hat{e}_u, \boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0$. By the (4.71c) and the homogeneous boundary condition in (4.82), we have $\langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \mathbf{P}_\partial^p\phi\rangle_{\partial\mathcal{T}_h} = 0$. Inserting these two zero terms into (4.86) and using the fact $\boldsymbol{\theta} = \nabla\phi$ leads to

$$\begin{aligned}\|e_u\|^2_{\mathcal{T}_h} = \ & (e_{\boldsymbol{q}} + \delta_{\boldsymbol{q}}, \quad \boldsymbol{\Pi}_h^p\boldsymbol{\theta} - \boldsymbol{\theta})_{\mathcal{T}_h} + \langle e_u - \hat{e}_u, \quad (\boldsymbol{\theta} - \boldsymbol{\Pi}_h^p\boldsymbol{\theta})\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} \\ & - \langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \quad \mathbf{P}_\partial^p\phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} + \langle e_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad \phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} \\ & + (\delta_{\boldsymbol{q}}, \nabla\phi)_{\mathcal{T}_h}.\end{aligned} \tag{4.87}$$

By the identity (4.72) and orthogonal properties in (4.69), we have

$$\begin{aligned}& \langle\boldsymbol{q}\cdot\boldsymbol{n} - \hat{\boldsymbol{q}}_h^p\cdot\boldsymbol{n}, \quad \mathbf{P}_\partial^p\phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} \\ = \ & \langle e_{\boldsymbol{q}}\cdot\boldsymbol{n} + \tau\tfrac{p}{h}(e_u - \hat{e}_u) + \tau\tfrac{p}{h}\mathbf{P}_\partial^p\delta_u, \quad \mathbf{P}_\partial^p\phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} \\ & + \langle\delta_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad \mathbf{P}_\partial^p\phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} \\ = \ & \langle e_{\boldsymbol{q}}\cdot\boldsymbol{n} + \tau\tfrac{p}{h}(e_u - \hat{e}_u) + \tau\tfrac{p}{h}\mathbf{P}_\partial^p\delta_u, \quad \phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h} \\ & + \langle\delta_{\boldsymbol{q}}\cdot\boldsymbol{n}, \quad \mathbf{P}_\partial^p\phi - \Pi_h^p\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned} \tag{4.88}$$

Then, substituting (4.88) into (4.87) and rearranging the resulted terms leads to the conclusion. □

Theorem 4 *Assume that (4.83) holds and the exact solution in (3.21) satisfies* $u|_K \in H^{s+1}(K) \cap L^2_{J^{-1}}(K)$ *and* $\boldsymbol{q}|_K \in (H^s(K))^2 \cap \left(L^2_{J^{-1}}(K)\right)^2$ *for any* $K \in \mathcal{T}_h$. *Then we have*

$$\|u - u_h^p\|_{\mathcal{T}_h} \le Ch^{s+1}p^{-s-1}(|u|_{s+1,\mathcal{T}_h} + |\boldsymbol{q}|_{s,\mathcal{T}_h}), \quad 1 \le s \le p. \tag{4.89}$$

Proof We are going to control each term in the expression given by Lemma 10. The following estimates are due to Theorems 1-3, Lemma 8 and the Cauchy-Schwarz's inequality. The first two terms can be estimated as

$$\begin{aligned}(e_{\boldsymbol{q}}+\delta_{\boldsymbol{q}},\quad \boldsymbol{\Pi}_h^p\boldsymbol{\theta}-\boldsymbol{\theta})_{\mathcal{T}_h} &\le (\|e_{\boldsymbol{q}}\|_{\mathcal{T}_h}+\|\delta_{\boldsymbol{q}}|_{\mathcal{T}_h})\|\boldsymbol{\Pi}_h^p\boldsymbol{\theta}-\boldsymbol{\theta}\|_{\mathcal{T}_h}\\ &\le Ch^{s+1}p^{-s-1}(|u|_{s+1,\Omega}+|\boldsymbol{q}|_{s,\mathcal{T}_h})|\boldsymbol{\theta}|_{1,\mathcal{T}_h},\\ \left\langle e_u-\hat{e}_u,\quad (\boldsymbol{\theta}-\boldsymbol{\Pi}_{\boldsymbol{h}}^{\boldsymbol{p}}\boldsymbol{\theta})\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} &\le Ch^{\frac{1}{2}}p^{-\frac{1}{2}}|||(e_u,e_{\boldsymbol{q}},\hat{e}_u)|||\cdot\|(\boldsymbol{\theta}-\boldsymbol{\Pi}_{\boldsymbol{h}}^{\boldsymbol{p}}\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_h}\\ &\le Ch^{s+1}p^{-s-1}(|u|_{s+1,\mathcal{T}_h}+|\boldsymbol{q}|_{s,\mathcal{T}_h})|\boldsymbol{\theta}|_{1,\mathcal{T}_h}.\end{aligned}$$

The estimates for the next three terms are given by

$$\begin{aligned}\left\langle\delta_{\boldsymbol{q}}\cdot\boldsymbol{n},\quad \mathbf{P}_\partial^p\phi-\Pi_h^p\phi\right\rangle_{\partial\mathcal{T}_h} &\le \|\delta_{\boldsymbol{q}}\|_{\partial\mathcal{T}_h}(\|\mathbf{P}_\partial^p\phi-\phi\|_{\partial\mathcal{T}_h}+\|\phi-\Pi_h^p\phi\|_{\partial\mathcal{T}_h})\\ &\le Ch^{s+\frac{3}{2}}p^{-s-1}(|u|_{s+1,\mathcal{T}_h}+|\boldsymbol{q}|_{s,\mathcal{T}_h})|\phi|_{2,\mathcal{T}_h},\\ \left\langle\tau\tfrac{p}{h}(e_u-\hat{e}_u),\quad \phi-\Pi_h^p\phi\right\rangle_{\partial\mathcal{T}_h} &\le \|\tau\tfrac{p}{h}(e_u-\hat{e}_u)\|_{\partial\mathcal{T}_h}\cdot\|\phi-\Pi_h^p\phi\|_{\partial\mathcal{T}_h}\\ &\le Ch^{-\frac{1}{2}}p^{\frac{1}{2}}|||e_u,e_{\boldsymbol{q}},\hat{e}_u|||\cdot\|\phi-\Pi_h^p\phi\|_{\partial\mathcal{T}_h}\\ &\le Ch^{s+1}p^{-s-1}(|u|_{s+1,\mathcal{T}_h}+|\boldsymbol{q}|_{s,\mathcal{T}_h})|\phi|_{2,\mathcal{T}_h},\end{aligned}$$

and

$$\left\langle\tau\tfrac{p}{h}\delta_u^p,\quad \phi-\Pi_h^p\phi\right\rangle_{\partial\mathcal{T}_h}\le C\tfrac{p}{h}\|\delta_u^p\|_{\partial\mathcal{T}_h}\|\phi-\Pi_h^p\phi\|_{\partial\mathcal{T}_h}\le Ch^{s+1}p^{-s-1}|u|_{s,\mathcal{T}_h}|\phi|_{2,\mathcal{T}_h}.$$

For the last term, let $\phi_p\in\mathcal{P}_p(K)\subset V_h^p(K)$ be the approximation of ϕ which can be estimates as in Lemma 2. By Proposition 2, we have $(\delta_{\boldsymbol{q}},\nabla\phi_p)_{\mathcal{T}_h}=0$. Therefore, using Lemma 2 leads to

$$(\delta_{\boldsymbol{q}},\nabla\phi)_{\mathcal{T}_h}=(\delta_{\boldsymbol{q}},\nabla\phi-\nabla\phi_p)_{\mathcal{T}_h}\le Ch^{s+1}p^{-s-1}|\boldsymbol{q}|_{s,\mathcal{T}_h}|\phi|_{2,\mathcal{T}_h}. \tag{4.90}$$

The above estimates together with the regularity assumption (4.83) imply

$$\|e_u\|_{\mathcal{T}_h}\le Ch^{s+1}p^{-s-1}(|\boldsymbol{q}|_{s,\mathcal{T}_h}+|u|_{s+1,\mathcal{T}_h}),\quad 1\le s\le p. \tag{4.91}$$

Then, the error estimate for u_h^p is obtained by applying triangular inequality. □

5 Numerical results

Now, we present some numerical results to verify the high-order convergence of the proposed HDTSEM. We show the convergence rates in both polynomial degree p and mesh size h for the elliptic problem on complex domains with non-smooth solutions.

In all numerical experiments, we test the elliptic problem (3.20) on various domains with $\beta=e^{x+y}$, $\gamma=1$. In the HDTSEM, we always set the penalty constant $\tau=1$, and choose the smooth exact solution:

$$u(x,y)=\cos(\pi(x^2+y^2)), \tag{5.92}$$

and non-smooth exact solutions:

$$u(x,y)=(x+y)^{\frac{5}{2}}, \tag{5.93a}$$

$$u(x,y)=(x-y)^{\frac{8}{3}}(e^{xy}+1). \tag{5.93b}$$

Hereafter, we denote by E_p the L^2-error of the numerical solution of u on a fixed mesh and with the polynomial degree p, E_h the L^2-error for a fixed polynomial

degree and mesh size h. All L^2-errors are calculated by using much higher order Gauss quadrature element-by-element. The convergence rates against p are measured by

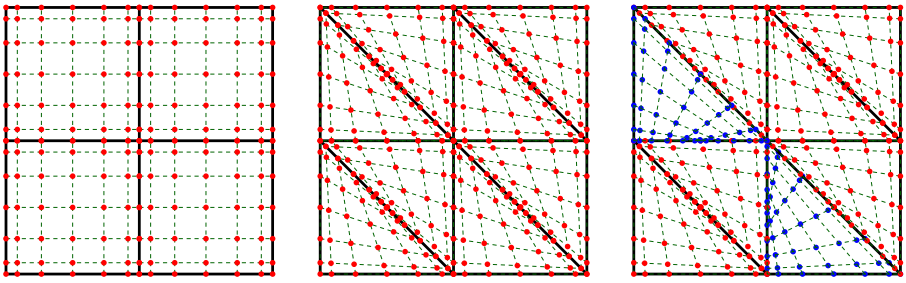
$$-\frac{\ln E_{p_k} - \ln E_{p_{k+1}}}{p_k - p_{k+1}}, \quad -\frac{\ln E_{p_k} - \ln E_{p_{k+1}}}{\ln p_k - \ln p_{k+1}},$$

which are the constants c and r in the expected convergence rates $\mathcal{O}(e^{-cp})$ and $\mathcal{O}(p^{-r})$ for smooth and non-smooth solutions, respectively. The convergence rate in h is gauged by

$$\frac{\ln E_{h_k} - \ln E_{h_{k+1}}}{\ln h_k - \ln h_{k+1}}.$$

Example 1 We first test the accuracy of the proposed HDTSEM by compared with hybridization discontinuous spectral element method (HDSEM) on Cartesian mesh. For this purpose, we set $\Omega = [0, 1]^2$ and the triangular meshes are generated by subdividing each element in the corresponding Cartesian meshes into two triangles, see Fig. 6 for the initial meshes and corresponding LGL-spectral element nodes distributions. As shown in our previous work [37], we can make the spectral element nodes generated on triangular meshes matching across the interior edges by applying the triangle-to-rectangle mapping appropriately to each triangular element (see Fig. 6 (b)). However, spectral element grids with unmatched nodes is inevitable if a similar cube-to-tetrahedron mapping is adopted for establishing HDTSEM for 3D problems on tetrahedral meshes. The HDTSEM proposed herein allows for unmatched nodes (see Fig. 6 (c)).

Therefore, we also compare the numerical results obtained by HDTSEM on spectral element grids with matched and unmatched nodes. In this case, solution (5.93a) and (5.93b) belong to $H^{4-\alpha}$ and $H^{3-\beta}$ for small $\alpha, \beta > 0$, respectively. Convergence rates presented in Tables 1 and 2 show that the proposed HDTSEM on triangular meshes and the HDSEM on Cartesian meshes share a very similar convergence behavior no matter spectral element grids with matched or unmatched nodes are used. Spectral accuracy is obtained for smooth solution and optimal convergence rate is observed for non-smooth solutions. Note that the non-smooth solutions (5.93a)–(5.93b) have boundary point and interior line singularities, respectively, in



(a) Cartesian mesh (b) Triangular mesh with matched nodes (c) Triangular mesh with unmatched nodes

Fig. 6 Initial Cartesian and triangular meshes and LGL-spectral element nodes distribution

Table 1 L^2-errors and convergence rates against p (fixed mesh) for (3.20)

	p	HDSEM		HDTSEM			
				Matched nodes		Unmatched nodes	
		Error	Rate	Error	Rate	Error	Rate
Smooth solution (5.92)	6	1.162E-04		1.744E-05		2.173E-05	
	9	2.255E-07	2.082	3.134E-08	2.107	4.665E-08	2.048
	12	2.772E-10	2.234	4.108E-11	2.212	5.333E-11	2.258
	15	2.458E-13	2.343	4.216E-14	2.294	1.833E-13	1.891
Non-smooth	6	5.180E-07		4.871E-07		4.879E-07	
solution (5.93a)	9	3.926E-08	6.363	3.808E-08	6.286	3.821E-08	6.282
	12	6.242E-09	6.392	6.079E-09	6.377	6.108E-09	6.373
	15	1.493E-09	6.411	1.447E-09	6.434	1.455E-09	6.429
Non-smooth	6	7.249E-05		5.920E-05		5.939E-05	
solution (5.93b)	9	1.802E-05	3.433	1.626E-05	3.187	1.627E-05	3.194
	12	7.031E-06	3.271	6.593E-06	3.138	6.595E-06	3.139
	15	3.432E-06	3.215	3.280E-06	3.129	3.281E-06	3.129

the computational domain. The convergence against p could be better than $\mathcal{O}(p^{-s})$ if a delicate weighted Sobolev norm is used for the convergence analysis, see [14,

Table 2 L^2-errors and convergence rates against h (fixed $p=6$) for elliptic problems on the domain $[0, 1]^2$

	h	HDSEM		HDTSEM			
				Matched nodes		Unmatched nodes	
		Error	Rate	Error	Rate	Error	Rate
Smooth solution	1/2	1.162E-04		1.744E-05		2.173E-05	
	1/4	8.487E-07	7.097	1.608E-07	6.761	1.930E-07	6.815
	1/8	4.646E-09	7.513	1.252E-09	7.005	1.546E-09	6.964
	1/16	2.830E-11	7.359	9.516E-12	7.039	1.201E-11	7.009
Non-smooth	1/2	5.180E-07		4.871E-07		4.879E-07	
solution (5.93a)	1/4	4.627E-08	3.485	4.347E-08	3.486	4.351E-08	3.487
	1/8	4.117E-09	3.490	3.866E-09	3.491	3.868E-09	3.492
	1/16	3.653E-10	3.494	3.429E-10	3.495	3.430E-10	3.495
Non-smooth	1/2	7.249E-05		5.920E-05		5.939E-05	
solution (5.93b)	1/4	8.019E-06	3.176	6.635E-06	3.158	6.646E-06	3.160
	1/8	8.926E-07	3.167	7.404E-07	3.164	7.411E-07	3.165
	1/16	9.939E-08	3.167	8.249E-08	3.166	8.255E-08	3.166

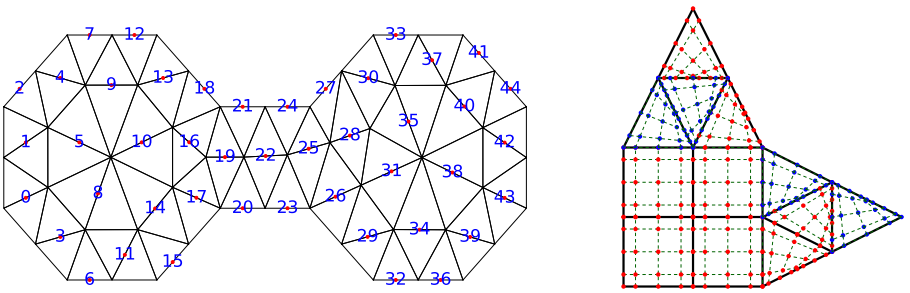


Fig. 7 Polygonal domains with triangular and hybrid meshes. (a) Triangular mesh. (b) Hybrid mesh

27] for more details. That is why better convergence rates against p are observed in Table 1 for non-smooth solutions.

Example 2 This example is to show the feasibility and accuracy of the proposed HDTSEM on unstructured meshes. Consider the polygonal domains with vertices given by

$$\begin{aligned}
&\text{polygon A:} \quad R\Big(\cos\theta_k + \cos\tfrac{\pi}{8},\ \sin\theta_k + \sin\tfrac{\pi}{8}\Big),\\
&\qquad\qquad R\Big(\cos\theta_k + 3\cos\tfrac{\pi}{8} + 2\sin\tfrac{\pi}{8},\ \sin\theta_k + \sin\tfrac{\pi}{8}\Big), \quad k = 0, 1, \cdots, 7,\\
&\text{polygon B:} \quad (0,0), (1,0), \Big(2, \tfrac{1}{2}\Big), (1,1), \Big(1, \tfrac{1}{2}\Big), (0,1)
\end{aligned}$$

Table 3 L^2-errors and convergence rates against p for fixed mesh (unstructured triangular mesh)

	p	Matched nodes		Unmatched nodes	
		Error	Rate	Error	Rate
Smooth solution	6	7.913E-04		4.901E-04	
	9	4.972E-06	1.690	3.609E-06	1.637
	12	9.787E-09	2.077	8.145E-09	2.031
	15	4.560E-12	2.557	5.878E-12	2.411
Non-smooth solution (5.93a)	6	6.754E-07		6.886E-07	
	9	1.217E-07	4.226	1.304E-07	4.105
	12	3.556E-08	4.278	3.956E-08	4.145
	15	1.352E-08	4.333	1.552E-08	4.194
Non-smooth solution (5.93b)	15	1.451E-06		1.263E-06	
	18	8.269E-07	3.083	7.324E-07	2.989
	21	5.196E-07	3.014	4.690E-07	2.891
	24	3.450E-07	3.068	3.118E-07	3.059

Table 4 L^2-errors and convergence rates against h for $p = 6$ (unstructured triangular mesh)

	h	Matched nodes		Unmatched nodes	
		Error	Rate	Error	Rate
Smooth solution	1/4	7.913E-04		4.901E-04	
	1/8	6.679E-06	6.888	6.066E-06	6.336
	1/16	4.406E-08	7.244	4.063E-08	7.222
	1/32	3.368E-10	7.032	3.117E-10	7.026
Non-smooth	1/4	6.754E-07		6.886E-07	
solution (5.93a)	1/8	8.554E-08	2.981	9.131E-08	2.915
	1/16	1.076E-08	2.991	1.145E-08	2.995
	1/32	1.350E-09	2.994	1.408E-09	3.024
Non-smooth	1/4	2.695E-05		2.360E-05	
solution (5.93b)	1/8	2.750E-06	3.293	2.720E-06	3.117
	1/16	3.013E-07	3.190	2.900E-07	3.230
	1/32	3.809E-08	2.984	3.617E-08	3.003

where $\theta_k = \frac{2k+1}{8}\pi$, $R = \frac{3}{4}\cos\frac{\pi}{8}$. Polygon A is triangulated by unstructured triangular meshes and polygon B is triangulated by hybrid meshes (with both rectangular and triangular elements), see Fig. 7 for initial meshes and spectral element nodes distributions. Solution (5.93a) and (5.93b) also have point and line singularity in the computational domain Ω and belong to $H^{4-\alpha}$ and $H^{3-\beta}$. For smooth solution

Table 5 L^2-errors and convergence rates against h left and against p right (mixed mesh)

	h	Error	Rate	p	Error	Rate
Smooth solution	1/2	5.847E-04		6	5.847E-04	
	1/4	1.061E-05	5.784	9	3.863E-06	1.673
	1/8	6.235E-08	7.411	12	2.213E-08	1.721
	1/16	4.549E-10	7.099	15	6.012E-11	1.969
Non-smooth	1/2	5.181E-07		6	5.181E-07	
solution (5.93a)	1/4	4.627E-08	3.485	9	3.927E-08	6.363
	1/8	4.117E-09	3.490	12	6.247E-09	6.390
	1/16	3.653E-10	3.494	15	1.496E-09	6.406
Non-smooth	1/2	7.281E-05		6	7.281E-05	
solution (5.93b)	1/4	8.033E-06	3.180	9	1.802E-05	3.445
	1/8	8.929E-07	3.169	12	7.031E-06	3.271
	1/16	9.940E-08	3.167	15	3.432E-06	3.214

(5.92), the L^2-errors of the HDTSEM using spectral element grids with matched or unmatched nodes decay like $\mathcal{O}(e^{-cp})$ for fixed unstructured mesh, i.e., spectral accuracy is obtained (see Tables 3 and 5). Again, optimal convergence rates against h are obtained, see Tables 4 and 5. For non-smooth solutions (5.93a)–(5.93b), better convergence rates are observed due to the same reason mentioned in Example 1.

6 Concluding remarks

In this paper, we have proposed a discontinuous triangular spectral element method on unstructured mesh using nodal basis. It is based on the hybridizable discontinuous formulation and enjoys the tensorial product property, flexibility in handling complex domains and significant reduction of global degree of freedoms. We have obtained optimal hp-error estimates in the L^2-norm with delicate treatment of the weak singularity induced by the triangle-to-rectangle transform. The numerical results have verified the expected convergence behaviors. We shall report the implementation and analysis of the 3D HDTSEM on unstructured tetrahedral meshes in a future work.

Acknowledgements The first and second authors received financial support provided by NSFC (grant 11771137, 12022104) and the Construct Program of the Key Discipline in Hunan Province. The research of the first author is partially supported by Hunan Provincial Innovation Foundation for Postgraduate (grant CX20190337). The research of the third author is supported by the Ministry of Education, Singapore, under its MOE AcRF Tier 2 Grants (MOE2018-T2-1-059 and MOE2017-T2-2-144). The fourth author is partially supported by NSFC (11771138).

Data availability The datasets supporting the conclusions of this article are included within the article.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Adams, R.A., Fournier, J.: Sobolev Spaces. Academic Press, New York (1975)
2. Canuto, C., Hussaini, M.Y., Quarteroni, A., Zang, T.A.: Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics. Scientific Computation Springer, Berlin (2007)
3. Chen, L., Shen, J., Xu, C.: A unstructured nodal spectral-element method for the Navier-Stokes equations. Commum. Comput. Phys. **12**, 315–336 (2012)
4. Chen, Q., Babuška, I.: Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. Comput. Methods Appl. Math. Eng. **128**, 405–417 (1995)
5. Ciarlet, P.G.: The finite element method for elliptic problems SIAM (2002)
6. 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**, 1319–1365 (2009)
7. Cockburn, B., Qiu, W., Shi, K.: Conditions for superconvergence of HDG methods for second-order elliptic problems. Math. Comp. **81**, 1327–1353 (2012)
8. Cockburn, B., Fu, G., Sayas, F.J.: Superconvergence by M-decompositions. Part i: General theory for HDG methods for diffusion. Math. Comp. **86**, 1609–1641 (2017)

9. Deville, M.O., Fischer, P.F., Fischer, P.F., Mund, E., et al.: High-Order Methods for Incompressible Fluid Flow, vol. 9. Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge (2002)
10. Dolejší, V., Feistauer, M.: Discontinuous Galerkin method: Analysis and applications to compressible flow. Springer Series in Computational Mathematics, p. 48 (2015)
11. Dubiner, M.: Spectral methods on triangles and other domains. J. Sci. Comput. **6**, 345–390 (1991)
12. Duffy, M.G.: Quadrature over a pyramid or cube of integrands with a singularity at a vertex. SIAM J. Numer. Anal. **19**, 1260–1262 (1982)
13. Egger, H., Waluga, C.: Hp analysis of a hybrid DG method for stokes flow. SIAM J. Numer. Anal. **33**, 687–721 (2013)
14. Georgoulis, E.H., Hall, E., Melenk, J.M.: On the suboptimality of the p-version interior penalty discontinuous Galerkiny method. J. Sci. Comput. **42**, 54–67 (2010)
15. Haupt, L., Stiller, J., Nagel, W.E.: A fast spectral element solver combining static condensation and multigrid techniques. J. Comput. Phys. **255**, 384–395 (2013)
16. Hesthaven, J.S.: From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. SIAM J. Numer. Anal. **35**, 655–676 (1998)
17. Hesthaven, J.S., Warburton, T.: Nodal discontinuous Galerkin methods: algorithms, analysis, and applications, vol. 54, Texts in Applied Mathematics Springer (2008)
18. Houston, P., Schwab, C., Süli, E.: Discontinuous hp-finite element methods for advection-diffusion-reaction problems. SIAM J. Numer. Anal. **39**, 2133–2163 (2002)
19. Huismann, I., Stiller, J., Fröhlich, J.: Factorizing the factorization–a spectral-element solver for elliptic equations with linear operation count. J. Comput. Phys. **346**, 437–448 (2017)
20. Karniadakis, G., Sherwin, S.: Spectral hp element methods for computational fluid dynamics, Numerical Mathematics and Scientific Computation, Oxford University Press New York (2005)
21. Kirby, R.M., Sherwin, S.J., Cockburn, B.: To CG or to HDG: a comparative study. J. Sci. Comput. **51**, 183–212 (2012)
22. Koornwinder, T.: Two-variable analogues of the classical orthogonal polynomials, in Theory and application of special functions. Elsevier, pp 435–495 (1975)
23. Kopriva, D.A.: Implementing spectral methods for partial differential equations: Algorithms for scientists and engineers, Scientific Computation Springer (2009)
24. Li, H., Wang, L.-L.: A spectral method on tetrahedra using rational basis functions. Int. J. Numer. Anal. Model. **7**, 330–355 (2010)
25. Li, J., Ma, H., Wang, L.-L., Wu, H.: Spectral element methods on hybrid triangular and quadrilateral meshes. Int. J. Numer. Anal. Model. **15**, 111–133 (2018)
26. Li, Y., Wang, L.-L., Li, H., Ma, H.: A new spectral method on triangles. Lecture Notes in Computational Sciences and Engineering **76**, 237–246 (2011)
27. Liu, W., Wang, L.-L., Wu, B.: Optimal error estimates for Legendre approximation of singular functions with limited regularity, arXiv:2006.00667 (2020)
28. Pasquetti, R., Rapetti, F.: Spectral element methods on unstructured meshes: which interpolation points? Numer. Algorithms **55**, 349–366 (2010)
29. Patera, A.T.: A spectral element method for fluid dynamics: laminar flow in a channel expansion. J. Comput. Phys. **54**, 468–488 (1984)
30. Pozrikidis, C.: Introduction to finite and spectral element methods using MATLAB Chapman & hall/CRC (2005)
31. Qiu, W., Shi, K.: An HDG method for convection diffusion equation. J. Sci. Comput. **66**, 346–357 (2016)
32. Samson, M.D., Li, H., Wang, L.-L.: A new triangular spectral element method i: implementation and analysis on a triangle. Numer. Algorithms **64**, 519–547 (2013)
33. Shan, W., Li, H.: The triangular spectral element method for Stokes eigenvalues. Math. Comp. **86**, 2579–2611 (2017)
34. Shen, J., Wang, L.-L., Li, H.: A triangular spectral element method using fully tensorial rational basis functions. SIAM J. Numer. Anal. **47**, 1619–1650 (2009)
35. Sherwin, S.J., Karniadakis, G.E.: A new triangular and tetrahedral basis for high-order (hp) finite element methods. Int. J. Numer. Meth. Eng. **38**, 3775–3802 (1995)
36. Taylor, M.A., Wingate, B.A., Vincent, R.E.: An algorithm for computing Fekete points in the triangle. SIAM J. Numer. Anal. **38**, 1707–1720 (2000)

37. Zhou, B., Wang, B., Wang, L.-L., Xie, Z.: A new triangular spectral element method II: Mixed formulation and hp-error estimates. Numer. Math. Theor. Meth. Appl. **12**, 72–97 (2019)