# A hybridizable direct discontinuous Galerkin method for elliptic problems 

Huiqiang Yue* © ${ }^{\text {© }}$, Jian Cheng, Tiegang Liu and Vladimir Shaydurov

"Correspondence:
yuehq@buaa.edu.cn
Key Laboratory of Mathematics, Informatics and Behavioral Semantics, Ministry of Education, School of Mathematics and Systems Science, Beihang University, Xueyuan Road No. 37, Haidian District, Beijing, China


#### Abstract

The aim of this work is to develop a hybridizable discontinuous Galerkin method for elliptic problems. In the proposed method, the numerical flux functions are constructed from the weak formulation of primal equation directly without converting the second-order equation to a first-order system. In order to guarantee the stability and convergence of the method, we derive a computable lower bound for the constant in numerical flux functions. We also establish a prior error estimation and give some theoretical analysis for the proposed method. Finally, a numerical experiment is presented to verify the theoretical results.


Keywords: hybridizable method; discontinuous Galerkin; elliptic problem

## 1 Introduction

In recent years, the discontinuous Galerkin method (DGM) has been extensively studied by lots of researchers on various problems since it was first introduced by Reed and Hill [1]. Particularly, for elliptic problems, a unified analysis framework for the most of DGMs proposed during the last 30 years has been given by Arnold et al. [2]. Unlike the continuous finite element method (FEM), DGM is able to easily handle non-conforming meshes and work well with varying polynomials on different cells. Therefore, DGM is much flexible for hp-adaptive approaches. For more details of DGM, we refer the reader to [3-6]. However, compared with continuous FEM, DGM results in a much larger linear system which leads to more computational costs. In order to overcome this issue, hybridizable discontinuous Galerkin method (HDGM) was proposed by Cockburn et al. in [7]. In HDGM, additional unknowns defined on interior faces are introduced and the global linear system only involves degrees of freedoms (DOFs) on interior faces. Therefore, this linear system is smaller than that obtained via classical DGM. On the other hand, recently, a new kind of DGM called direct discontinuous Galerkin method (DDGM) was introduced by Liu and Yan $[8,9]$. One of the key aspects of DDGM is that it directly uses the weak formulation of the original equation instead of rewriting the equations into a first-order system. In this way, the DDGM is easy to implement and more effective because of there being no extra computation for auxiliary variables.

The purpose of this work is to develop a hybridizable discontinuous Galerkin method, which combines the advantages of HDGM and DDGM together, based on the main idea of DDGM. The rest of the paper is organized as follows. In Section 2, we introduce the model problem and some basic notations which will be used in this work. Then the full
derivation of our hybridizable direct discontinuous Galerkin method is given in Section 3. In Section 4, we discuss the theoretical error analysis and derive computable lower bounds for the constant coefficients in numerical fluxes in order to guarantee the stability and convergence of this method. In Section 5, a numerical example is presented to verify the theoretical analysis results. Finally, in Section 6, we give some concluding remarks.

## 2 Model problem and basic notations

In this work, for the sake of simplicity, we consider the following elliptic boundary value problem:

$$
\begin{align*}
& -\Delta u=f \quad \text { in } \Omega  \tag{1}\\
& u=g \quad \text { on } \partial \Omega \tag{2}
\end{align*}
$$

where $\Omega$ is a bounded polygonal domain in $\mathbb{R}^{2}$ and $f \in L^{2}(\Omega)$.
Let $\mathcal{T}_{h}=\{\kappa\}$ be the subdivision of domain $\Omega$. For two adjacent elements $\kappa^{+}$and $\kappa^{-}, \partial k^{+} \cap$ $\partial k^{-}$is the interior face. For $\kappa \in \mathcal{T}_{h}, \partial k \cap \partial \Omega$ is the boundary face. We denote by $\mathcal{E}_{h}^{I}$ the set of interior faces and $\mathcal{E}_{h}^{B}$ the set of boundary faces. Also, we denote all the faces of $\mathcal{T}_{h}$ by $\mathcal{E}_{h}:=\mathcal{E}_{h}^{I} \cup \mathcal{E}_{h}^{B}$. In addition, we denote the outward unit normals of $\partial \kappa^{+}$and $\partial \kappa^{-}$by $\mathbf{n}^{+}$and $\mathbf{n}^{-}$, respectively. For $\mathcal{E} \in \mathcal{E}_{h}^{I}$, we define the jumps

$$
[\mathbf{Q} \cdot \mathbf{n}]_{\mathcal{E}}=\mathbf{Q}^{+} \mathbf{n}^{+}+\mathbf{Q}^{-} \mathbf{n}^{-},
$$

where $\mathbf{Q}$ is a vector valued function and $(\cdot)^{ \pm}$are the interior traces on $\partial k^{ \pm}$, respectively. In order to approximate the modal problem using discontinuous finite elements, some useful spaces are defined as follows:

$$
\begin{aligned}
& H^{s}\left(\mathcal{T}_{h}\right):=\left\{u \in L^{2}(\Omega):\left.u\right|_{\kappa} \in H^{s}(\kappa), \forall \kappa \in \mathcal{T}_{h}\right\}, \\
& \mathcal{V}:=\left\{u: u \in H^{2}\left(\mathcal{T}_{h}\right) \cap H^{1}(\Omega)\right\}, \\
& \mathcal{V}_{h}:=\left\{u \in L^{2}(\Omega):\left.u\right|_{\kappa} \in \mathcal{P}^{p}(\kappa), \forall \kappa \in \mathcal{T}_{h}\right\},
\end{aligned}
$$

and the traced finite element space

$$
\begin{aligned}
& \mathcal{M}:=\left\{\hat{u}: \hat{u} \in L^{2}\left(\mathcal{E}_{h}\right)\right\}, \\
& \mathcal{M}_{h}:=\left\{\hat{u} \in L^{2}\left(\mathcal{E}_{h}\right):\left.\hat{u}\right|_{\mathcal{E}} \in \mathcal{P}^{p}(\mathcal{E}), \forall \mathcal{E} \in \mathcal{E}_{h}\right\} .
\end{aligned}
$$

Here $\mathcal{P}^{p}$ is the polynomial of degree at most $p$. In addition, we introduce the space with Dirichlet boundary conditions

$$
\mathcal{M}_{h}(g)=\left\{u \in \mathcal{M}_{h}: u=g \text { on } \partial \Omega\right\} .
$$

## 3 The hybridizable direct discontinuous Galerkin method

### 3.1 Solving the local problems

In order to derive hybridizable discontinuous Galerkin method, we restrict the modal problem to any element $\kappa \in \mathcal{T}_{h}$ not the entire computational domain. Therefore, if we
know the trace value $\hat{u}$ on $\partial \kappa$, we can solve the following local problem to obtain $u$ inside $\kappa$ :

$$
\begin{align*}
& -\Delta u=f \quad \text { in } \kappa,  \tag{3}\\
& u=\hat{u} \quad \text { on } \partial \kappa . \tag{4}
\end{align*}
$$

Thus, we can multiply the local problem (3) by a test function $v_{h}$ and integrate by parts to obtain the following variational problem over $\kappa$ : find $u_{h} \in \mathcal{V}_{h}$ such that

$$
\begin{equation*}
\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}-\left(\nabla u_{h} \cdot \mathbf{n}, v_{h}\right)_{\partial \kappa}=\left(f, v_{h}\right)_{\kappa}, \quad \forall v_{h} \in \mathcal{V}_{h}, \tag{5}
\end{equation*}
$$

where $\mathbf{n}$ is the unit outward normal to $\partial \kappa$. In two dimensional case, based on the main idea of direct discontinuous Galerkin method [8, 9], we approximate $\nabla u_{h}:=\left(u_{x}, u_{y}\right)$ via numerical flux function $\widehat{\nabla u_{h}}=\left(\widehat{u_{x}}, \widehat{u_{y}}\right)$, i.e.,

$$
\begin{align*}
& \widehat{u_{x}}=\beta \frac{\hat{u}-u_{h}}{h_{\kappa}} n_{1}+u_{x},  \tag{6}\\
& \widehat{u_{y}}=\beta \frac{\hat{u}-u_{h}}{h_{\kappa}} n_{2}+u_{y},
\end{align*}
$$

where $n_{i}$ is the $i$ th component of $\mathbf{n}, \beta$ is a constant coefficient and $h_{\kappa}$ is the characteristic length of $\partial k$. In this work, $h_{\kappa}$ is chosen as the diameter of the cell and for two dimensional quadrilateral mesh $h_{\kappa}$ is approximated by the length of longest diagonal. For more complicated cases, we refer the reader to [10]. Therefore, we get the discretization of the local problem: find $u_{h} \in \mathcal{V}_{h}$ such that

$$
\begin{equation*}
\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}-\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, v_{h}\right)_{\partial \kappa}=\left(f, v_{h}\right)_{\kappa}, \quad \forall v_{h} \in \mathcal{V}_{h}, \tag{7}
\end{equation*}
$$

where

$$
\begin{equation*}
\widehat{\nabla u_{h}} \cdot \mathbf{n}=\beta \frac{\hat{u}-u_{h}}{h_{\kappa}}+\nabla u_{h} \cdot \mathbf{n} . \tag{8}
\end{equation*}
$$

### 3.2 Defining the global problem

In the above derivation, we have given the discretization of the local problem when the trace value $\hat{u}$ is known. However, in general, the value of $\hat{u}$ is unknown. Therefore, we need additional conditions to overcome this issue. In this work, we weakly enforce the flux is normal-continuous, which is the essential condition of hybridizable discontinuous Galerkin method. Thus, the function $\hat{u}$ can be computed as the solution of the following problems:

$$
\begin{align*}
& \left(\left[\widehat{\nabla u_{h}} \cdot \mathbf{n}\right], \hat{v}_{h}\right)_{\mathcal{E}}=0 \quad \text { if } \mathcal{E} \in \mathcal{E}_{h}^{I}, \forall \hat{v}_{h} \in \mathcal{M}_{h},  \tag{9}\\
& \hat{u}=g \quad \text { if } \mathcal{E} \in \mathcal{E}_{h}^{B} \tag{10}
\end{align*}
$$

Summing (7) and (9) over all $\kappa \in \mathcal{T}_{h}$, we have

$$
\begin{equation*}
\sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}-\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, v_{h}\right)_{\partial \kappa}\right\}=\sum_{\kappa \in \mathcal{T}_{h}}\left(f, v_{h}\right)_{\kappa}, \quad \forall v_{h} \in \mathcal{V}_{h} \tag{11}
\end{equation*}
$$

$$
\begin{equation*}
\sum_{\mathcal{E} \in \mathcal{E}_{h}^{I}}\left(\left[\widehat{\nabla u_{h}} \cdot \mathbf{n}\right], \hat{v}_{h}\right)_{\mathcal{E}}=0, \quad \forall \hat{v}_{h} \in \mathcal{M}_{h} \tag{12}
\end{equation*}
$$

We note that (12) is a face-based form, i.e., each interior face occurs only once. In the following, we rewrite this face-based form to element-based form, namely

$$
\begin{equation*}
\sum_{\kappa \in \mathcal{T}_{h}}\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, \hat{v}_{h}\right)_{\partial \kappa \backslash \mathcal{E}_{h}^{B}}=\sum_{\mathcal{E} \in \mathcal{E}_{h}^{I}}\left(\left[\widehat{\nabla u_{h}} \cdot \mathbf{n}\right], \hat{v}_{h}\right)_{\mathcal{E}}=0 \tag{13}
\end{equation*}
$$

Then we add (13) to (11) and enforce the boundary conditions, from which one derives a new discrete form: find $u_{h} \in \mathcal{V}_{h}$ and $\hat{u}_{h} \in \mathcal{M}(g)$ such that for all $v_{h} \in \mathcal{V}_{h}$ and $\hat{v}_{h} \in \mathcal{M}(0)$ we have

$$
\begin{align*}
& \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}+\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa \backslash \mathcal{E}_{h}^{B}}\right\} \\
& \quad+\sum_{\kappa \in \mathcal{T}_{h}}\left(\frac{\beta}{h_{\kappa}}\left(g-u_{h}\right)+\nabla u_{h} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa \in \mathcal{E}_{h}^{B}}=\sum_{\kappa \in \mathcal{T}_{h}}\left(f, v_{h}\right)_{\kappa} . \tag{14}
\end{align*}
$$

Moreover, we rewrite the above equation into equivalent form, i.e., find $u_{h} \in \mathcal{V}_{h}$ and $\hat{u}_{h} \in$ $\mathcal{M}(0)$ such that for all $v_{h} \in \mathcal{V}_{h}$ and $\hat{v}_{h} \in \mathcal{M}(0)$ we have

$$
\begin{gather*}
\sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}+\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa}\right\} \\
=\sum_{\kappa \in \mathcal{T}_{h}}\left(f, v_{h}\right)_{\kappa}+\sum_{\kappa \in \mathcal{T}_{h}} \frac{\beta}{h_{\kappa}}\left(g, v_{h}\right)_{\partial \kappa \in \mathcal{E}_{h}^{B}} . \tag{15}
\end{gather*}
$$

Note that (15) is asymmetric. Thus, it is hard to guarantee the optimal $L^{2}$ error convergence. In order to overcome this problem, additional interface integral terms are added to the left hand side of (15) to make the discretization symmetric. For more introduction to the symmetric direct discontinuous Galerkin method, $c f$. [11]. The final discretization is given as follows: find $u_{h} \in \mathcal{V}_{h}$ and $\hat{u}_{h} \in \mathcal{M}_{h}(0)$ such that for all $v_{h} \in \mathcal{V}_{h}$ and $\hat{v}_{h} \in \mathcal{M}_{h}(0)$ we have

$$
\begin{equation*}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right)=\mathcal{L}\left(\mathbf{v}_{h}\right), \tag{16}
\end{equation*}
$$

where

$$
\begin{align*}
& \mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right):=\sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}+\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa}+\left(\widehat{\nabla v_{h}} \cdot \mathbf{n}, \hat{u}_{h}-u_{h}\right)_{\partial \kappa}\right\},  \tag{17}\\
& \mathcal{L}\left(\mathbf{v}_{h}\right):=\sum_{\kappa \in \mathcal{T}_{h}}\left(f, v_{h}\right)_{\kappa}+\sum_{\kappa \in \mathcal{T}_{h}} 2 \frac{\beta}{h_{\kappa}}\left(g, v_{h}\right)_{\partial \kappa \in \mathcal{E}_{h}^{B}}-\sum_{\kappa \in \mathcal{T}_{h}}\left(\nabla v_{h} \cdot \mathbf{n}, g\right)_{\partial \kappa \in \mathcal{E}_{h}^{B}} \tag{18}
\end{align*}
$$

and

$$
\begin{aligned}
& \mathbf{u}_{h}:=\left(u_{h}, \hat{u}_{h}\right), \quad \mathbf{v}_{h}:=\left(v_{h}, \hat{v}_{h}\right), \\
& \widehat{\nabla u_{h}} \cdot \mathbf{n}=\beta \frac{\hat{u}_{h}-u_{h}}{h_{\kappa}}+\nabla u_{h} \cdot \mathbf{n}
\end{aligned}
$$

$$
\widehat{\nabla v_{h}} \cdot \mathbf{n}=\beta \frac{\hat{v}_{h}-v_{h}}{h_{\kappa}}+\nabla v_{h} \cdot \mathbf{n} .
$$

Remark 1 Although the form of (16) is similar to the hybridizable symmetric interior penalty method (HSIP) [12], the derivation and motivation of the two approaches are different. For the discretization (16), the numerical flux is directly approximated by the solution gradient. For HSIP method, the jumps are introduced as penalty terms to guarantee the stability of the method.

### 3.3 Hybridization

Utilizing traces defined on element faces, we can eliminate DoFs related with $u_{h}$ to obtain a global linear system only involving DoFs of traces through static condensation process, as described below.

After assembling all of the cell and face terms in (16) together, we have the following linear system:

$$
\left(\begin{array}{ll}
A^{\left(u_{h}, v_{h}\right)} & B^{\left(\hat{u}_{h}, v_{h}\right)}  \tag{19}\\
C^{\left(u_{h}, \hat{v}_{h}\right)} & D^{\left(\hat{u}_{h}, v_{h}\right)}
\end{array}\right)\binom{U^{\left(u_{h}\right)}}{\hat{U}^{\left(\hat{u}_{h}\right)}}=\binom{F^{\left(v_{h}\right)}}{G^{\left(\hat{v}_{h}\right)}} .
$$

Here we divide the global matrix into four blocks $A^{\left(u_{h}, v_{h}\right)}, B^{\left(\hat{u}_{h}, v_{h}\right)}, C^{\left(u_{h}, \hat{v}_{h}\right)}, D^{\left(\hat{u}_{h}, \hat{v}_{h}\right)}$. The superscript on the block matrix or vector means how the values of this matrix are coupling. For example, $A^{\left(u_{h}, v_{h}\right)}$ means this block matrix only consists of terms only involving the integrals containing both of $u_{h}$ and $v_{h}$. Additionally, $U^{\left(u_{h}\right)}$ and $\hat{U}^{\left(\hat{u}_{h}\right)}$ represent the vector of unknowns for $u_{h}$ and $\hat{u}_{h}$, respectively.

By the block matrix operations, we obtain

$$
\begin{align*}
& A U+B \hat{U}=F  \tag{20}\\
& C U+D \hat{U}=G \tag{21}
\end{align*}
$$

where, in order to simplify the formulas, we drop the superscripts of block matrices and vectors. Thus, $U$ can be eliminated by some algebraic operations, i.e.,

$$
\begin{equation*}
\left(D-C A^{-1} B\right) \hat{U}=G-C A^{-1} F \tag{22}
\end{equation*}
$$

Therefore, we obtain a reduced global linear system only for $\hat{U}$. After solving the system (22), we can get $U$ immediately through an element-by-element fashion, namely

$$
\begin{equation*}
U=A^{-1} F-A^{-1} B \hat{U} . \tag{23}
\end{equation*}
$$

Remark 2 We note that in (22) and (23) the inversion of $A$ is needed. However, because of the fact that DoFs related with $u_{h}$ are not coupled between different elements, $A$ is easy to invert.

## 4 A prior error estimation

In this section we give a prior error estimation of the method proposed above and derive computable lower bounds for the constant $\beta$ in numerical flux functions in order to
guarantee the stability and convergence of this method. Before giving the theoretical error analysis, let us define the following two energy norms associated to bilinear form (17):

$$
\begin{align*}
& \|\mathbf{u}\|_{*}^{2}:=\sum_{\kappa \in \mathcal{T}_{h}}\left\{\|\nabla u\|_{\kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}\right\}  \tag{24}\\
& \|\mathbf{u}\|^{2}:=\sum_{\kappa \in \mathcal{T}_{h}}\left\{\|\nabla u\|_{\kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}+h_{\kappa}\|\nabla u \cdot \mathbf{n}\|_{\partial \kappa}^{2}\right\} \tag{25}
\end{align*}
$$

where $\|\cdot\|$ is the $L^{2}$ norm. Moreover, let us recall the following $L^{2}$-projection error estimation inequality [13].

Lemma 1 Let $\mathcal{P}_{h}$ be the $L^{2}$-projection onto $\mathcal{V}_{h}$ and $u \in H^{p+1}(\kappa), 0 \leq m \leq p$, then

$$
\begin{align*}
& \left\|u-\mathcal{P}_{h} u\right\|_{H^{m}(\kappa)} \leq C h_{\kappa}^{p+1-m}|u|_{H^{(p+1)}(\kappa)}  \tag{26}\\
& \left\|\left.u\right|_{\partial \kappa}-\left.\mathcal{P}_{h} u\right|_{\partial \kappa}\right\|_{\partial \kappa} \leq C h_{\kappa}^{p+\frac{1}{2}}|u|_{H^{(p+1)}(\kappa)} \tag{27}
\end{align*}
$$

where $\kappa \in \mathcal{T}_{h}$.

Theorem 1 Let the bilinear form $\mathcal{B}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right)$ be defined by (17), then we have:
(i) If $u$ is the exact solution of (1) and $\hat{u}$ is the trace value of $u$. Then

$$
\begin{equation*}
\mathcal{B}_{h}(\mathbf{u}, \mathbf{v})=\mathcal{L}(\mathbf{v}), \quad \forall v \in \mathcal{V}, \hat{v} \in \mathcal{M}(0) \tag{28}
\end{equation*}
$$

where $\mathbf{u}:=(u, \hat{u})$ and $\mathbf{v}:=(v, \hat{v})$. Furthermore, if $u_{h} \in \mathcal{V}_{h}$ and $\hat{u}_{h} \in \mathcal{M}_{h}(0)$, then

$$
\begin{equation*}
\mathcal{B}_{h}\left(\mathbf{u}_{h}-\mathbf{u}, \mathbf{v}_{h}\right)=0, \quad \forall v \in \mathcal{V}_{h}, \hat{v}_{h} \in \mathcal{M}_{h}(0) . \tag{29}
\end{equation*}
$$

(ii) Assume that $\theta_{\kappa}$ is the smallest angle in the two triangles obtained by subdividing the quadrilateral cell $\kappa \in \mathcal{T}_{h}$ along the longest diagonal and $\gamma_{\kappa}$ is a positive constant associated to $\kappa$. For all $\kappa \in \mathcal{T}_{h}$, define

$$
\beta_{\kappa}^{*}=\frac{p^{\kappa}\left(p^{\kappa}+1\right)}{\gamma_{\kappa} \sin \theta_{\kappa}}
$$

where $p^{\kappa}$ is the degree of the polynomial defined on $\kappa$. For all $u_{h} \in \mathcal{V}_{h}, \hat{u}_{h} \in \mathcal{M}_{h}$, if $\gamma_{\kappa}<1$ and $\beta>\beta_{\kappa}^{*}$, then there are constants $C$, $\alpha$ independent of mesh size, such that

$$
\begin{equation*}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{u}_{h}\right) \geq C\| \| \mathbf{u}_{h}\left\|_{*}^{2} \geq \frac{C}{1+\alpha}\right\| \mathbf{u}_{h} \|^{2} \tag{30}
\end{equation*}
$$

where $C=\min _{\kappa \in \mathcal{T}_{h}}\left\{\min \left\{1-\gamma_{\kappa}, 2\left(\beta-\beta_{\kappa}^{*}\right)\right\}\right\}$ and $\alpha=\max _{\kappa \in \mathcal{T}_{h}}\left\{\frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{\sin \theta_{\kappa}}\right\}$.
(iii) For all $u, v \in \mathcal{V}$ and $\hat{u}, \hat{v} \in \mathcal{M}$, we have

$$
\begin{equation*}
\left|\mathcal{B}_{h}(\mathbf{u}, \mathbf{v})\right| \leq C\|\mathbf{u}\|\| \| \mathbf{v} \|, \tag{31}
\end{equation*}
$$

where $C=1+2 \beta$.

Proof First of all, we rewrite the bilinear form $\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right)$ as follows:

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right)= & \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}+\left(\widehat{\nabla u_{h}} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa}+\left(\widehat{\nabla v_{h}} \cdot \mathbf{n}, \hat{u}_{h}-u_{h}\right)_{\partial \kappa}\right\} \\
= & \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\nabla u_{h}, \nabla v_{h}\right)_{\kappa}+2 \frac{\beta}{h_{\kappa}}\left(\hat{u}_{h}-u_{h}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa}\right. \\
& \left.+\left(\nabla u_{h} \cdot \mathbf{n}, \hat{v}_{h}-v_{h}\right)_{\partial \kappa}+\left(\nabla v_{h} \cdot \mathbf{n}, \hat{u}_{h}-u_{h}\right)_{\partial \kappa}\right\} .
\end{aligned}
$$

(i) Here $u$ is the exact solution, thus,

$$
\left.(u-\hat{u})\right|_{\mathcal{E} \in \mathcal{E}_{h}^{I}}=0,
$$

therefore,

$$
\begin{aligned}
\mathcal{B}_{h}(\mathbf{u}, \mathbf{v})= & \sum_{\kappa \in \mathcal{T}_{h}}\left\{(\nabla u, \nabla v)_{\kappa}+(\nabla u \cdot \mathbf{n}, \hat{v}-v)_{\partial \kappa}\right\} \\
= & \sum_{\kappa \in \mathcal{T}_{h}}\left\{-(\Delta u, v)+(\nabla u \cdot \mathbf{n}, \hat{v})_{\partial \kappa}\right\}+\sum_{\mathcal{E} \in \mathcal{E}_{h}^{B}} 2 \frac{\beta}{h_{\kappa}}(-g,-v)_{\mathcal{E}}+\sum_{\mathcal{E} \in \mathcal{E}_{h}^{B}}(\nabla v \cdot \mathbf{n},-g)_{\mathcal{E}} \\
= & \sum_{\kappa \in \mathcal{T}_{h}}(f, v)_{\kappa}+\underbrace{\sum_{\mathcal{E} \in \mathcal{E}_{h}^{I}}(\nabla u \cdot \mathbf{n}, \hat{v})_{\mathcal{E}}}_{=0}+\underbrace{\sum_{=0}(\nabla u \cdot \mathbf{n}, \hat{v})_{\mathcal{E}}}_{\mathcal{E} \in \mathcal{E}_{h}^{B}} \\
& +\sum_{\mathcal{E} \in \mathcal{E}_{h}^{B}} 2 \frac{\beta}{h_{\kappa}}(g, v)_{\mathcal{E}}-\sum_{\mathcal{E} \in \mathcal{E}_{h}^{B}}(\nabla v \cdot \mathbf{n}, g)_{\mathcal{E}} \\
= & \mathcal{L}(\mathbf{v}) .
\end{aligned}
$$

Furthermore, due to $v_{h} \in \mathcal{V}, \hat{v}_{h} \in \mathcal{M}(0)$, we have

$$
\begin{equation*}
\mathcal{B}_{h}\left(\mathbf{u}, \mathbf{v}_{h}\right)=\mathcal{L}\left(\mathbf{v}_{h}\right), \tag{32}
\end{equation*}
$$

therefore,

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}-\mathbf{u}, \mathbf{v}_{h}\right) & =\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{v}_{h}\right)-\mathcal{B}_{h}\left(\mathbf{u}, \mathbf{v}_{h}\right) \\
& =\mathcal{L}\left(\mathbf{v}_{h}\right)-\mathcal{L}\left(\mathbf{v}_{h}\right) \\
& =0 .
\end{aligned}
$$

(ii) For all $u_{h} \in \mathcal{V}_{h}, \hat{u}_{h} \in \mathcal{M}_{h}$, employing the Cauchy-Schwarz inequality we have

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{u}_{h}\right) & =\sum_{\kappa \in \mathcal{T}_{h}}\left\{\|\nabla u\|_{\kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2}+2\left(\nabla u_{h} \cdot \mathbf{n}, \hat{u}_{h}-u_{h}\right)_{\partial \kappa}\right\} \\
& \geq \sum_{\kappa \in \mathcal{T}_{h}}\left\{\|\nabla u\|_{\kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2}-2\left\|\nabla u_{h} \cdot \mathbf{n}\right\|_{\partial \kappa}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}\right\} .
\end{aligned}
$$

Figure 1 The subdivision of quadrilateral mesh.


Then we recall the inverse trace inequality for the two dimensional simplex [14]:

$$
\begin{equation*}
\left\|u_{h}\right\|_{\partial \kappa} \leq \sqrt{\frac{\left(p^{\kappa}+1\right)\left(p^{\kappa}+2\right)}{2} \frac{|\partial \kappa|}{|\kappa|}}\left\|u_{h}\right\|_{\kappa}, \quad \forall u_{h} \in \mathcal{P}^{p^{\kappa}}(\kappa) . \tag{33}
\end{equation*}
$$

For a quadrilateral cell with edges $\mathcal{E}_{1}, \mathcal{E}_{2}, \mathcal{E}_{3}, \mathcal{E}_{4}$, we subdivide it into two triangles $T_{1}, T_{2}$ along the longest diagonal, as shown in Figure 1, and the angle between $\mathcal{E}_{i}$ and the longest diagonal is denoted by $\theta_{i}, i=1,2,3,4$. Then we have

$$
\begin{aligned}
& |\partial \kappa|=\sum_{i=1}^{4}\left|\mathcal{E}_{i}\right|, \\
& \left|T_{1}\right|=\frac{1}{2} h_{\kappa}\left|\mathcal{E}_{1}\right| \sin \theta_{1}=\frac{1}{2} h_{\kappa}\left|\mathcal{E}_{2}\right| \sin \theta_{2} .
\end{aligned}
$$

Therefore,

$$
\left|T_{1}\right|=\frac{1}{4} h_{\kappa}\left(\left|\mathcal{E}_{1}\right| \sin \theta_{1}+\left|\mathcal{E}_{2}\right| \sin \theta_{2}\right)
$$

Similarly,

$$
\left|T_{2}\right|=\frac{1}{4} h_{\kappa}\left(\left|\mathcal{E}_{3}\right| \sin \theta_{3}+\left|\mathcal{E}_{4}\right| \sin \theta_{4}\right)
$$

Thus,

$$
|\kappa|=\left|T_{1}\right|+\left|T_{2}\right|=\frac{1}{4} h_{\kappa} \sum_{i=1}^{4}\left(\left|\mathcal{E}_{i}\right| \sin \theta_{i}\right) \text {. }
$$

Here we assume $\theta_{\kappa}$ is the smallest angle of $\left\{\theta_{i}\right\}$, namely, $\theta_{\kappa}=\min \left\{\theta_{i}\right\}$, then

$$
|\kappa| \geq \frac{1}{4} h_{\kappa} \sin \theta_{\kappa} \sum_{i=1}^{4}\left|\mathcal{E}_{i}\right|
$$

Finally, we obtain

$$
\frac{|\partial \kappa|}{|\kappa|}=\frac{4}{h_{\kappa}} \frac{\sum_{i=1}^{4}\left|\mathcal{E}_{i}\right|}{\sum_{i=1}^{4}\left(\left|\mathcal{E}_{i}\right| \sin \theta_{i}\right)} \leq \frac{4}{h_{\kappa} \sin \theta_{\kappa}} .
$$

Therefore, applying the inverse trace inequality (33), we have

$$
\begin{equation*}
\left\|\nabla u_{h} \cdot \mathbf{n}\right\|_{\partial \kappa} \leq\left\|\nabla u_{h}\right\|_{\partial \kappa} \leq \sqrt{\frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{h_{\kappa} \sin \theta_{\kappa}}}\left\|\nabla u_{h}\right\|_{\kappa} . \tag{34}
\end{equation*}
$$

Furthermore, employing Young's inequality, we have

$$
\begin{aligned}
\left\|\nabla u_{h} \cdot \mathbf{n}\right\|_{\partial \kappa}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa} & \leq \sqrt{\frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{h_{\kappa} \sin \theta_{\kappa}}}\left\|\nabla u_{h}\right\|_{\kappa}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa} \\
& \leq \frac{\gamma_{\kappa}}{2}\left\|\nabla u_{h}\right\|_{\kappa}^{2}+\frac{1}{2 \gamma_{\kappa}} \frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{h_{\kappa} \sin \theta_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2},
\end{aligned}
$$

where $\gamma_{\kappa}$ is a positive constant. Finally, we obtain

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{u}_{h}\right) & \geq \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left\|\nabla u_{h}\right\|_{\kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2}-2\left\|\nabla u_{h} \cdot \mathbf{n}\right\|_{\partial \kappa}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}\right\} \\
& \geq \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(1-\gamma_{\kappa}\right)\left\|\nabla u_{h}\right\|_{\kappa}^{2}+2 \frac{\beta-\frac{p^{\kappa}\left(p^{\kappa}+1\right)}{\gamma_{\kappa} \sin \theta_{\kappa}}}{h_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2}\right\} .
\end{aligned}
$$

Thus, for all $\kappa \in \mathcal{T}_{h}$, if $\gamma_{\kappa}<1$ and $\beta>\frac{p^{\kappa}\left(p^{\kappa}+1\right)}{\gamma_{\kappa} \sin \theta_{\kappa}}$, we have

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{u}_{h}\right) & \geq \sum_{\kappa \in \mathcal{T}_{h}}\left\{\min \left\{1-\gamma_{\kappa}, 2\left(\beta-\frac{p^{\kappa}\left(p^{\kappa}+1\right)}{\gamma_{\kappa} \sin \theta_{\kappa}}\right)\right\}\left(\left\|\nabla u_{h}\right\|_{\kappa}^{2}+\frac{1}{h_{\kappa}}\left\|\hat{u}_{h}-u_{h}\right\|_{\partial \kappa}^{2}\right)\right\} \\
& \geq C\left\|u_{h}\right\|_{*}^{2},
\end{aligned}
$$

where $C=\min _{\gamma \in \mathcal{T}_{h}}\left\{\min \left\{1-\gamma_{\kappa}, 2\left(\beta-\frac{p^{\kappa}\left(p^{\kappa}+1\right)}{\gamma_{\kappa} \sin \theta_{\kappa}}\right)\right\}\right\}$.
Moreover, applying inequality (34), we have

$$
\begin{aligned}
\left\|\mathbf{u}_{h}\right\|^{2}-\left\|\mathbf{u}_{h}\right\|_{*}^{2} & =\sum_{\kappa \in \mathcal{T}_{h}} h_{\kappa}\left\|\nabla u_{h} \cdot \mathbf{n}\right\|_{\partial \kappa}^{2} \\
& \leq \sum_{\kappa \in \mathcal{T}_{h}} \frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{\sin \theta_{\kappa}}\left\|\nabla u_{h}\right\|_{\kappa}^{2} \\
& \leq \max _{\kappa \in \mathcal{T}_{h}}\left\{\frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{\sin \theta_{\kappa}}\right\} \sum_{\kappa \in \mathcal{T}_{h}}\left\|\nabla u_{h}\right\|_{\kappa}^{2} \\
& \leq \alpha\left\|\mathbf{u}_{h}\right\|_{{ }_{*}}^{2},
\end{aligned}
$$

where $\alpha=\max _{\kappa \in \mathcal{T}_{h}}\left\{\frac{2 p^{\kappa}\left(p^{\kappa}+1\right)}{\sin \theta_{\kappa}}\right\}$.
Therefore,

$$
\left\|\mathbf{u}_{h}\right\|_{*}^{2} \geq \frac{1}{1+\alpha}\left\|\mathbf{u}_{h}\right\|^{2}
$$

Thus, we have

$$
\mathcal{B}_{h}\left(\mathbf{u}_{h}, \mathbf{u}_{h}\right) \geq C\left\|\mathbf{u}_{h}\right\|_{*}^{2} \geq \frac{C}{1+\alpha}\left\|\mathbf{u}_{h}\right\|^{2} .
$$

This completes the proof.
(iii) Applying the Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
\left|\mathcal{B}_{h}(\mathbf{u}, \mathbf{v})\right| \leq & \sum_{\kappa \in \mathcal{T}_{h}}\left\{\|\nabla u\|_{\kappa}\|\nabla v\|_{\kappa}+\|\nabla u \cdot \mathbf{n}\|_{\partial \kappa}\|\hat{v}-v\|_{\partial \kappa}\right. \\
& \left.+\|\nabla v \cdot \mathbf{n}\|_{\partial \kappa}\|\hat{u}-u\|_{\partial \kappa}+2 \frac{\beta}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}\|\hat{v}-v\|_{\partial \kappa}\right\} \\
\leq & \sum_{\kappa \in \mathcal{T}_{h}}\left\{\left(\|\nabla u\|_{\kappa}^{2}+h_{\kappa}\|\nabla u \cdot \mathbf{n}\|_{\partial \kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}\right)^{\frac{1}{2}}\right. \\
& \left.\cdot\left(\|\nabla v\|_{\kappa}^{2}+h_{\kappa}\|\nabla v \cdot \mathbf{n}\|_{\partial \kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{v}-v\|_{\partial \kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\|\hat{v}-v\|_{\partial \kappa}^{2}\right)^{\frac{1}{2}}\right\} \\
\leq & \left(\sum_{\kappa \in \mathcal{T}_{h}}\left(\|\nabla u\|_{\kappa}^{2}+h_{\kappa}\|\nabla u \cdot \mathbf{n}\|_{\partial \kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}\right)\right)^{\frac{1}{2}} \\
& \cdot\left(\sum_{\kappa \in \mathcal{T}_{h}}\left(\|\nabla v\|_{\kappa}^{2}+h_{\kappa}\|\nabla v \cdot \mathbf{n}\|_{\partial \kappa}^{2}+\frac{1}{h_{\kappa}}\|\hat{v}-v\|_{\partial \kappa}^{2}+2 \frac{\beta}{h_{\kappa}}\|\hat{v}-v\|_{\partial \kappa}^{2}\right)\right)^{\frac{1}{2}} \\
= & \left(\|\mathbf{u}\|^{2}+\sum_{\kappa \in \mathcal{T}_{h}} 2 \frac{\beta}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2}\right)^{\frac{1}{2}} \cdot\left(\|\mathbf{v}\|^{2}+\sum_{\kappa \in \mathcal{T}_{h}} 2 \frac{\beta}{h_{\kappa}}\|\hat{v}-v\|_{\partial \kappa}^{2}\right)^{\frac{1}{2}} .
\end{aligned}
$$

Based on the definition of the energy norm, we have

$$
\sum_{\kappa \in \mathcal{T}_{h}} \frac{1}{h_{\kappa}}\|\hat{u}-u\|_{\partial \kappa}^{2} \leq\|\mathbf{u}\|^{2}
$$

Therefore,

$$
\begin{aligned}
\left|\mathcal{B}_{h}(\mathbf{u}, \mathbf{v})\right| & \leq\left(\|\mathbf{u}\|\left\|^{2}+2 \beta\right\| \mathbf{u} \|^{2}\right)^{\frac{1}{2}} \cdot\left(\| \| \mathbf{v}\left\|^{2}+2 \beta\right\| \mathbf{v} \|^{2}\right)^{\frac{1}{2}} \\
& =C\|\mathbf{u}\|\| \| \mathbf{v} \|,
\end{aligned}
$$

where $C=1+2 \beta$.
This completes the proof.
Corollary 1 Assume that $\theta_{\min }=\min _{\kappa \in \mathcal{T}_{h}}\left\{\theta_{\kappa}\right\}$ and the uniform polynomials are used in the computational domain. Then the threshold value of $\beta$ is obtained when $\gamma_{\kappa}=1$, i.e., $\beta_{\mathrm{ths}}=$ $\frac{p(p+1)}{\sin \theta_{\min }}$.

Theorem 2 Let $u \in H^{(p+1)}\left(\mathcal{T}_{h}\right)$ be the exact solution of (1) and $u_{h} \in \mathcal{V}_{h}$ be the solution of (16). Then there is constant $C>0$ such that

$$
\begin{equation*}
\left\|\mathbf{u}-\mathbf{u}_{h}\right\| \| \leq C h^{p}|u|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)}, \tag{35}
\end{equation*}
$$

where $h=\max _{\kappa \in \mathcal{T}_{h}} h_{\kappa}$ and $C$ is independent of $h$.
Proof Employing (30), we have

$$
\mathcal{B}_{h}\left(\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}, \mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right) \geq C_{0}\left\|\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right\|^{2}
$$

where $\mathcal{P}_{h} \mathbf{u}:=\left(\mathcal{P}_{h} u, \mathcal{P}_{h} \hat{u}\right)$. Furthermore,

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}, \mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right) & \stackrel{(29)}{=} \mathcal{B}_{h}\left(\mathbf{u}-\mathcal{P}_{h} \mathbf{u}, \mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right) \\
& \stackrel{(31)}{\leq} C_{1}\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\|\| \| \mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\| \| .
\end{aligned}
$$

Therefore, we have

$$
C_{0}\left\|\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right\|\left\|^{2} \leq C_{1}\right\|\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\|\| \| \mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u} \|
$$

namely,

$$
\left\|\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right\| \left\lvert\, \leq \frac{C_{1}}{C_{0}}\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\|\right.
$$

Then

$$
\begin{aligned}
\left\|\mathbf{u}-\mathbf{u}_{h}\right\| \| & \leq\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\| \mid+\left\|\mathbf{u}_{h}-\mathcal{P}_{h} \mathbf{u}\right\| \\
& \leq\left(1+\frac{C_{1}}{C_{0}}\right)\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\| .
\end{aligned}
$$

Using Lemma 1, we can deduce

$$
\left\|\mathbf{u}-\mathcal{P}_{h} \mathbf{u}\right\| \leq \leq C_{2} h^{p}|u|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)} .
$$

Finally, we obtain

$$
\left.\left\|\mathbf{u}-\mathbf{u}_{h}\right\|\left|\leq C h^{p}\right| u\right|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)} .
$$

Theorem 3 Let $u \in H^{(p+1)}\left(\mathcal{T}_{h}\right)$ be the exact solution of (1) and $u_{h} \in \mathcal{V}_{h}$ be the solution of (16). Then there is a constant $C>0$ such that

$$
\begin{equation*}
\left\|\mathbf{u}-\mathbf{u}_{h}\right\|_{L^{2}\left(\mathcal{T}_{h}\right)} \leq C h^{p+1}|u|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)}, \tag{36}
\end{equation*}
$$

where $h=\max _{\kappa \in \mathcal{T}_{h}} h_{\kappa}$ and $C$ is independent of $h$.

Proof We note that $\mathcal{B}_{h}(\cdot, \cdot)$ is symmetric and discretization (16) is consistent (Theorem 1). Thus, the discretization (16) is adjoint consistent. Therefore, we can define the following adjoint problem:

Find $z \in \mathcal{V}, \hat{z} \in \mathcal{M}(0), \forall v \in \mathcal{V}, \hat{v} \in \mathcal{M}(0)$ such that

$$
\mathcal{B}_{h}(\mathbf{v}, \mathbf{z})=\left(u-u_{h}, v\right)_{\Omega} .
$$

Particularly, we take $\mathbf{v}=\mathbf{u}-\mathbf{u}_{h}$, then

$$
\left\|u-u_{h}\right\|_{L^{2}\left(\mathcal{T}_{h}\right)}^{2}=\mathcal{B}_{h}\left(\mathbf{u}-\mathbf{u}_{h}, \mathbf{z}\right) .
$$

For all $z_{h} \in \mathcal{V}_{h}, \hat{z}_{h} \in \mathcal{M}_{h}(0)$, by employing (29), we obtain

$$
\begin{aligned}
\mathcal{B}_{h}\left(\mathbf{u}-\mathbf{u}_{h}, \mathbf{z}\right) & =\mathcal{B}_{h}\left(\mathbf{u}-\mathbf{u}_{h}, \mathbf{z}-\mathbf{z}_{h}\right) \\
& \stackrel{(31)}{\leq} C_{1}\| \| \mathbf{u}-\mathbf{u}_{h}\| \|\left\|\mathbf{z}-\mathbf{z}_{h}\right\| \| .
\end{aligned}
$$

From Theorem 2, we have

$$
\begin{aligned}
& \left.\left\|\mathbf{u}-\mathbf{u}_{h}\right\|\left|\leq C h^{p}\right| u\right|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)}, \\
& \left|\left\|\mathbf{z}-\mathbf{z}_{h}\right\|\right| \leq C h|z|_{H^{2}\left(\mathcal{T}_{h}\right)} .
\end{aligned}
$$

And we assume that the elliptic regularity holds [15], i.e.,

$$
|z|_{H^{2}} \leq C\left\|u-u_{h}\right\|_{L^{2}} .
$$

Finally, combining the above formulas together, we deduce

$$
\left\|\mathbf{u}-\mathbf{u}_{h}\right\|_{L^{2}\left(\mathcal{T}_{h}\right)} \leq C h^{p+1}|u|_{H^{(p+1)}\left(\mathcal{T}_{h}\right)},
$$

where $C$ is a constant and independent of $h$.

## 5 Numerical results

In this section, we present a numerical experiment to assess the practical performance of the proposed method and verify the theoretical analysis. The right hand side $f$ and boundary condition are chosen such that the analytical solution is $u(x, y)=\cos (8 \pi x)+$ $\cos (8 \pi y)$, and $\Omega:=[0,1]^{2}$. The analytical solution of this problem is shown in Figure 2. We carry out the numerical computation on a uniform mesh with 1,024 elements which is shown in Figure 3. For this uniform mesh $\theta_{\min }=\frac{\pi}{4}$. Therefore, the threshold value $\beta^{*}=$ $\sqrt{2} p(p+1)$. Figures 4 and 5 show the numerical solution under different $\beta$ for $p=1,2$, respectively. From these results we note that for a smaller $\beta$ the numerical solution is unstable, while for $\beta>\beta^{*}$ no oscillation occurs. Furthermore, we plot the error $\left\|u-u_{h}\right\|_{L_{2}}$ under different $\beta$ in Figures 6 and 7. From these figures, we again observe that for a smaller $\beta$ the oscillation occurs and when $\beta>\beta^{*}$ the $L_{2}$ error tends to be stable. Moreover, we can

Figure 2 Exact solution.


Figure 3 Structured mesh with 1,024 elements.


Figure 4 Numerical solution for $p=1$ under
$\boldsymbol{d i f f e r e n t} \boldsymbol{\beta} \boldsymbol{\beta} \boldsymbol{\beta}=1$ (top), $\beta=2$ (middle), $\beta=4$ (bottom).


Figure 5 Numerical solution for $\boldsymbol{p}=\mathbf{2}$ under different $\boldsymbol{\beta} . \boldsymbol{\beta}=1$ (top), $\boldsymbol{\beta}=5$ (middle), $\boldsymbol{\beta}=9$ (bottom).


Figure 6 The error $\left\|u-u_{h}\right\|_{L_{2}}$ under different $\beta$ on the same mesh for $p=1$.


Figure 7 The error $\left\|u-u_{h}\right\|_{L_{2}}$ under different $\beta$ on the same mesh for $p=2$.


Table 1 Convergence of the error $\left\|u-u_{h}\right\|_{L^{2}}$ with $p=1,2,3$

| Elements | $p=1$ |  | $p=2$ |  | $p=3$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Error | Order | Error | Order | Error | Order |
| 64 | $4.229 \mathrm{e}-01$ | - | $1.270 \mathrm{e}-01$ | - | 1.285e-02 | - |
| 256 | $9.523 \mathrm{e}-02$ | 2.15 | $1.211 \mathrm{e}-02$ | 3.39 | 1.285e-03 | 3.32 |
| 1,024 | $2.331 \mathrm{e}-02$ | 2.03 | $1.529 \mathrm{e}-03$ | 2.99 | $8.043 \mathrm{e}-05$ | 4.00 |
| 4,096 | 5.802e-03 | 2.01 | $1.916 \mathrm{e}-04$ | 3.00 | $5.030 \mathrm{e}-06$ | 4.00 |

also get the 'numerical lower bounds for $\beta$ ' that are equal to 3 for $p=1$ and 6 for $p=2$, which confirms the effectivity of the theoretical lower bounds of $\beta$. Table 1 shows the history of the convergence under global mesh refinement with a fixed $\beta>\beta^{*}$. From these results we note that the convergence rate is optimal in $L^{2}$ norm, i.e., for $p=1,2,3$, the order of convergence is $p+1$. This result coincides with the prior error analysis which has been shown in Theorem 3.

## 6 Conclusion

In this paper, we have proposed a hybridizable discontinuous Galerkin method based on DDGM. During the derivation of this method, we introduced the main idea of HDGM and showed how to construct numerical flux based on DDGM. Then a rigorous theoretical analysis has been established. We also gave computable lower bounds for the constant in the numerical flux functions. The numerical results obtained coincided with the theoretical analysis and the optimal convergence order has been observed. Although in this work our analysis is based on the Dirichlet boundary condition, we can easily extend this method to the Neumann or Robin boundary condition, which we leave here as future research work.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

The first author had the idea of this work and drafted the manuscript. The second and third authors gave helpful suggestions on the practical implementation of this method. The fourth author provided some useful theoretical analysis. All authors read and approved the final manuscript.

## Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 91530325), the International Cooperation Project (No. 2010DFR00700), and the Fundamental Research of Civil Aircraft (No. MJ-F-2012-04).

## References

1. Reed, WH, Hill, TR: Triangular mesh methods for the neutron transport equation. Los Alamos Report LA-UR-73-479 (1973)
2. Arnold, DN, Brezzi, F, Cockburn, B, Marini, LD: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. 39(5), 1749-1779 (2001)
3. Douglas, J, Dupont, T: Interior penalty procedures for elliptic and parabolic Galerkin methods. In: Computing Methods in Applied Sciences, pp. 207-216. Springer, New York (1976)
4. Cockburn, B, Hou, S, Shu, C-W: The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. Math. Comput. 54(190), 545-581 (1990)
5. Cockburn, B, Shu, C-W: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. Math. Comput. 52(186), 411-435 (1989)
6. Cockburn, B, Lin, S-Y, Shu, C-W: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. J. Comput. Phys. 84(1), 90-113 (1989)
7. Cockburn, B, Gopalakrishnan, J, Lazarov, R: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. 47(2), 1319-1365 (2009)
8. Liu, H, Yan, J: The direct discontinuous Galerkin (DDG) method for diffusion with interface corrections. Commun. Comput. Phys. 8(3), 541-564 (2010)
9. Liu, H, Yan, J: The direct discontinuous Galerkin (DDG) methods for diffusion problems. SIAM J. Numer. Anal. 47(1), 675-698 (2009)
10. Cheng, J, Yang, X, Liu, T, Luo, H: A direct discontinuous Galerkin method for the compressible Navier-Stokes equations on arbitrary grids. In: 54th AIAA Aerospace Sciences Meeting, p. 1334 (2016)
11. Vidden, C, Yan, J: A new direct discontinuous Galerkin method with symmetric structure for nonlinear diffusion equations. J. Comput. Math. 31(6), 638-662 (2013)
12. Lehrenfeld, C, Schöberl, J, Reusken, A: Hybrid discontinuous Galerkin methods for solving incompressible flow problems. PhD thesis, Rheinisch-Westfälischen Technischen Hochschule Aachen (2010)
13. Hartmann, R: Numerical analysis of higher order discontinuous Galerkin finite element methods (2008)
14. Warburton, T, Hesthaven, JS: On the constants in $h p$-finite element trace inverse inequalities. Comput. Methods Appl. Mech. Eng. 192(25), 2765-2773 (2003)
15. Brenner, S, Scott, R: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics. Springer, New York (2008)

## Submit your manuscript to a SpringerOpen ${ }^{\ominus}$ journal and benefit from:

Convenient online submission

- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online

High visibility within the field

- Retaining the copyright to your article

