DISCONTINUOUS GALERKIN METHOD FOR MONOTONE NONLINEAR ELLIPTIC PROBLEMS

CHUNJIA BI AND YANPING LIN

Abstract. In this paper, we consider the incomplete interior penalty method for a class of second order monotone nonlinear elliptic problems. Using the theory of monotone operators, we show that the corresponding discrete method has a unique solution. The a priori error estimate in an energy norm is developed under the minimal regularity assumption on the exact solution, i.e., \( u \in H^1(\Omega) \). Moreover, we propose a residual-based a posteriori error estimator and derive the computable upper and lower bounds on the error in an energy norm.

Key words. discontinuous Galerkin method, nonlinear elliptic problems, monotone, a priori error estimate, a posteriori error estimate.

1. Introduction

The discontinuous Galerkin (DG) methods were introduced in the early 1970s to solve first-order hyperbolic problems [17, 34, 38, 46]. Simultaneously, but quite independently, as non-standard schemes, they were proposed for the approximations of second-order elliptic equations [1, 41, 56]. Since the DG methods are locally conservative, stable and high-order methods, which can easily handle irregular meshes with hanging nodes and approximations that have polynomials of different degree in different elements, they have been studied extensively in the past several decades. We refer the reader to [2, 15, 16] for a comprehensive historical survey of this area of research, to [1, 11, 12, 23, 29, 42, 44, 45, 47, 50, 55, 56] and [52, 58] for the a priori error analysis of the DG methods for linear elliptic problems and optimal control problems, respectively.

Except for linear elliptic problems, some researchers have studied the a priori error estimates of the DG methods for the nonlinear elliptic problems. Houston, Robson and S"ul"u [30] considered a one parameter family of \( hp \)-DG methods for a class of quasi-linear elliptic problems with mixed boundary conditions

\[
-\nabla \cdot (\lambda(x, |\nabla u|) \nabla u) = f(x),
\]

where the function \( \lambda \) satisfies the following monotone condition, i.e., there exist positive constants \( m_\lambda \) and \( M_\lambda \) such that

\[
m_\lambda (t - s) \leq \lambda(x,t)t - \lambda(x,s)s \leq M_\lambda (t - s), \quad t \geq s \geq 0, \quad x \in \Omega.
\]

Using a result from the theory of monotone operators, the authors shown that the corresponding discrete method has a unique solution and derived the a priori error estimate in a mesh-dependent energy norm for \( u \in C^1(\Omega) \cap H^k(\Omega), k \geq 2 \), which is optimal in the mesh size and mildly suboptimal in the polynomial degree.

Received by the editors December 1, 2011 and, in revised form, June 22, 2012.

2000 Mathematics Subject Classification. 65N15; 65N30.

The research of C. Bi was partially supported by Shandong Province Natural Science Foundation (ZR2010AM004), Projects of Shandong Province Higher Educational Science and Technology Program (J09LA01, J10LA01, J11LA09), and the AMSS-PolyU Joint Research Institute for Engineering and Management Mathematics.
Gudi, Nataraj, and Pani [27], Gudi and Pani [28] studied the $hp$ local DG method and the $hp$-DG methods, respectively, for a class of quasilinear elliptic problems of nonmonotone type

\begin{equation}
-\nabla \cdot (a(x,u)\nabla u) = f(x), \quad \text{in } \Omega,
\end{equation}

proving the existence and uniqueness of the discrete solution and derived the a priori error estimates in a mesh-dependent energy norm and in the $L^2$-norm under the assumption $u \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$. Bi and Ginting [4] considered the two-grid algorithm of the $h$-version DG method for (1.2) and derived the convergence estimates.

Recently, Gudi, Nataraj and Pani [26] analyzed a one parameter family of $hp$-DG methods for the following second order nonlinear elliptic boundary value problems

\begin{equation}
-\nabla \cdot a(x,u,\nabla u) + a_0(x,u,\nabla u) = f(x), \quad \text{in } \Omega,
\end{equation}

where the given functions $a(x,y,z)$ and $a_0(x,y,z)$ are twice continuously differentiable with all the derivatives through second order being bounded, and the matrix $a_z(x,y,z)$ is symmetric and there exist two positive constants $\lambda_1$ and $\lambda_2$ such that

\begin{equation}
\lambda_1|\xi|^2 \leq \xi^T a_z(x,y,z)\xi \leq \lambda_2|\xi|^2, \quad \forall x \in \Omega, \quad \forall y \in \mathbb{R}, \quad \forall z, \xi \in \mathbb{R}^2.
\end{equation}

The authors developed the error estimate in the broken $H^1$-norm, which is optimal in $h$ and suboptimal in $p$, using piecewise polynomials of degree $p \geq 2$, when the solution $u \in H^{3/2}(\Omega)$. Additionally, we mention some related works in which the $h$-DG methods are used to solve the other nonlinear problems. We refer to [8] and [9] for (1.1) and monotone nonlinear fluid flow problems respectively, to [39] for nonlinear dispersive problems, to [43] for the nonlinear second-order elliptic and hyperbolic systems, to [48] for nonlinear non-Fickian diffusion problems and to [53] for nonlinear elasticity problems.

On the other hand, the a posteriori error estimates of DG methods have attracted many researchers’ attention and some important results have been achieved. For the linear elliptic problems, we refer the reader to [3, 13, 19, 31, 32, 35, 36, 49] and the references therein for details. However, there are considerably fewer papers that are concerned with the nonlinear elliptic problems. To the best of our knowledge, there are only [7] and [33] in this direction. Bustinza, Gatica and Cockburn [7] used a Helmholtz decomposition of the error to derive a residual-based a posteriori error estimates in an energy norm of $h$-version local DG method for the nonlinear elliptic problems (1.2) in which the differential operators are strongly monotone. Similar technique has been used in [3, 13] for linear elliptic problems. Houston, S"{u}li and Wihler [33] derived energy norm a posteriori error estimates of the $hp$-DG methods for (1.1) using the technique of the approximation of discontinuous finite element functions by conforming ones, which has been developed by some authors in the context of the $h$-DG methods in [35, 36, 37] and has been extended to $hp$-DG methods by [31, 33].

In this paper, we study the incomplete interior penalty method for the nonlinear elliptic problems that have the form (1.3) and are monotonic (specific assumptions on the functions $a_i$, $i = 0, 1, 2$, where $a = (a_1, a_2)$, will be given in subsection 2.1). Our purpose in this paper is twofold. As a first task, we formulate the incomplete interior penalty method to the monotone nonlinear elliptic problems and prove that the form associated with this DG method is bounded, Lipschitz-continuous and strongly monotone. Then, using a result from the theory of monotone operators,
we show that this DG method has a unique solution. Using the technique proposed
in [25], in which the a priori error estimates in the energy norm of various DG
methods have been developed under the assumption \( u \in H^k(\Omega) \) for the linear
elliptic problems of order \( 2k, k = 1, 2 \), we develop the a priori error estimate in
an energy norm under the minimal regularity assumption on the exact solution,
i.e., \( u \in H^1(\Omega) \). In contrast to [25], in this paper, we consider the DG method for
monotone nonlinear elliptic problems. A difficulty and a novel contribution is the
a posteriori error analysis.

The second task in this paper is to carry out the a posteriori error analysis. We
analyze the residual-based a posteriori error estimates of the incomplete interior
penalty method for the monotone nonlinear elliptic problems and derive the com-
putable upper bounds on the error in the broken \( H^1 \)-seminorm and in an energy
norm. The proof of our upper bound crucially relies on the approximation of discon-
tinuous finite element functions by conforming ones. In particular, by introducing
a function in \( H^1_0(\Omega) \) and its conforming approximation, we give a representa-
tion of the error, which plays a key role in the derivation of the upper bound for the
gradients of the error. Based on this representation of the error, with the help of
the approximation result, we show that our error estimator proposed in this paper
is reliable with respect to the broken \( H^1 \)-seminorm and an energy norm, respec-
tively. It should be pointed out that our upper bounds on the error in the broken
\( H^1 \)-seminorm and in an energy norm don’t contain the penalty parameter, which
appears in the analogous upper bound, see [35, 36] for Possion equation. In this
respect, our upper bounds on the error in the broken \( H^1 \)-seminorm and the energy
norm are stronger than those in [35, 36].

We remark that the DG method in this paper is the so-called incomplete interior
penalty method. This method was studied by [18, 43, 51]. Houston, Robson and
Süli [30], Gudi, Nataraj and Pani [26] analyzed a one parameter family \( hp \)-DG
method, in which the parameter \( \theta = 0 \) corresponds to the incomplete interior
penalty method.

We point out that the classical finite element method of the monotone nonlin-
ear elliptic problems considered in this paper has been studied in [21, 22, 57], in
which the solvability of the discrete problems and the convergence of approximates
solutions to an exact weak solution \( u \in H^1(\Omega) \) are proved.

The outline of this paper is as follows. In Section 2, we introduce the continuous
problems and formulate the incomplete interior penalty method. In Section 3,
we prove the existence and uniqueness of the DG solution. Since a result, which is
similar to the discrete local efficiency estimate in the a posteriori error analysis, will
be used to derive the a priori error estimate in the energy norm, we first discuss the
a posteriori error estimates in Section 4. And Section 5 is devoted to the a priori
error estimate in the energy norm.

Throughout this paper, we use the following standard notation. For simplicity, in
this paper, we assume that \( \Omega \) is a bounded polygonal domain in \( \mathbb{R}^2 \) with boundary
\( \partial \Omega \). For the domain \( \Omega \), we write \( W^{m,p}(\Omega) \), \( 1 \leq p \leq \infty \), to denote the usual Sobolev
spaces with norm \( \| \cdot \|_{m,p,\Omega} \) and seminorm \( | \cdot |_{m,p,\Omega} \) \cite{6, 14}. To simplify the notation,
we denote \( W^{m,2}(\Omega) \) by \( H^m(\Omega) \) and skip the index \( p = 2 \) and \( \Omega \) whenever possible,
i.e., we use \( |u|_{m,\Omega} = |u|_{m,p,\Omega} \), \( |u|_m,\Omega = |u|_{m,2,\Omega} \) and \( |u|_0 = |u| \). The same
convention is used for the seminorms as well. We define \( H^1_0(\Omega) \) to be the subspace
of \( H^1(\Omega) \) in which the functions have zero trace on \( \partial \Omega \). In what follows, the symbol
\( (\cdot, \cdot) \) denotes the \( L^2(\Omega) \) inner product. Moreover, \( C \), with or without subscripts,
denote positive constants independent of $h$ and may take different values at different occurrences.

2. Preliminaries

In this section, we first recall the continuous problems and introduce some assumptions on the coefficients functions in the subsection 2.1. The triangulation $T_h$ of the domain $\Omega$ and the discontinuous finite element space associated with $T_h$ are given in the subsection 2.2. In the subsection 2.3, we formulate the incomplete interior penalty method of the monotone nonlinear elliptic problems.

2.1. Continuous problems. In this subsection, we again recall the following monotone nonlinear elliptic boundary value problem

$$
-\nabla \cdot a(x, u, \nabla u) + a_0(x, u, \nabla u) = f(x), \quad \text{in } \Omega,
$$

$$
u = 0, \quad \text{on } \partial \Omega,
$$

where $a = (a_1, a_2)$.

We assume that the given functions $a_i(x, \xi), i = 0, 1, 2$, have the following properties:

(A). The derivatives $\frac{\partial a_i}{\partial \xi_j}(x, \xi), (i, j = 0, 1, 2)$ are continuous and bounded in $\Omega \times \mathbb{R}^3$, i.e., there exists a constant $C_0 > 0$ such that

$$
\left|\frac{\partial a_i}{\partial \xi_j}(x, \xi)\right| \leq C_0, \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^3.
$$

(B). The derivatives $\frac{\partial a_i}{\partial \xi_j}(x, \xi)(i, j = 0, 1, 2)$ satisfy the following inequality

$$
\alpha \sum_{i=1}^{2} \eta_i^2 \leq \sum_{i,j=0}^{2} \frac{\partial a_i}{\partial \xi_j}(x, \xi) \eta_i \eta_j, \quad \forall x \in \Omega, \quad \forall \xi, \eta \in \mathbb{R}^3,
$$

where $\alpha > 0$ is a constant independent of $x, \xi$ and $\eta$.

Remark 2.1. In contrast with the assumptions on the functions $a_i$ which are given in [26], in this paper, we have (B) instead of (1.4), which guarantees (2.1) is monotonic.

Examples of the functions $a_i(x, \xi)$. Let $b(x, t)$ be a function defined on $\Omega \times [0, +\infty)$ with the following properties:

a). $b(x, t)$ and the derivatives $\frac{\partial b}{\partial x_i}, i = 1, 2, \frac{\partial b}{\partial t}$ are continuous in $\Omega \times [0, +\infty)$.

b). There exist constants $0 < c_1 < c_2$ such that

$$
c_1 \leq b(x, t) \leq c_2, \quad \text{in } \Omega \times [0, +\infty),
$$

and

$$
\frac{\partial b}{\partial x_i}(x, t^2) \leq c_2, i = 1, 2; \quad 0 \leq \frac{\partial b}{\partial t} \leq c_2, \quad \forall x \in \Omega, t \geq 0,
$$

and

$$
\frac{\partial b}{\partial \tau}(x, \tau^2) |\tau| \leq c_2, \quad \frac{\partial b}{\partial \tau}(x, \tau^2) \tau^2 \leq c_2, \quad \forall x \in \Omega, \forall \tau \in \mathbb{R}.
$$

Let us set

$$
a_i(x, \xi) = b(x, \xi_0^2 + \xi_1^2 + \xi_2^2) \xi_i, \quad i = 0, 1, 2,
$$

or

$$
a_i(x, \xi) = b(x, \xi_1^2 + \xi_2^2) \xi_i, \quad i = 1, 2; \quad a_0 = \tilde{b}(x, \xi_0^2) \xi_0,
$$

(2.3)

or

(2.4)
where $b, \tilde{b}$ are two functions with the properties introduced above. Functions (2.3) and (2.4) satisfy the assumptions (A)-(B) and the problem (2.1) defined by means of (2.3) and (2.4) has many important applications (cf. e.g. [20, 24]). In particular, we note that the prescribed mean curvature presented in [26], also falls into the form (2.1) with our assumptions:

$$a(x, u, \nabla u) = (1 + |\nabla u|^2)^{-\frac{1}{2}} \nabla u, \quad a_0(x, u\nabla u) = 0.$$  

2.2. Discontinuous finite element space. In this paper, we consider shape-regular meshes $\mathcal{T}_h$ that partition $\Omega$ into open triangles or quadrilaterals, where $h = \max \{h_K : K \in \mathcal{T}_h\}$ and $h_K$ is the diameter of $K$. For a definition of shape regularity, we refer to [14]. Each element $K \in \mathcal{T}_h$ can be affinely mapped onto the reference element $\hat{K}$, which is either the open triangle $\{(x_1, x_2) : -1 < x_1 < 1, -1 < x_2 < -x_1\}$ or the open unit square $(-1,1)^2$ in $\mathbb{R}^2$.

In this paper, we do not require $\mathcal{T}_h$ to be conforming. In this case, we allow the mesh $\mathcal{T}_h$ to be one-irregular, i.e., each edge of any one element contains at most one hanging node (which, for simplicity, we assume to be the midpoint of the corresponding edge), see [33] for details.

Due to our assumption that the subdivision $\mathcal{T}_h$ is shape-regular, we know that it satisfies the bounded local variation condition, that is, if $|\partial K_i \cap \partial K_j| > 0$ for any $K_i, K_j \in \mathcal{T}_h$, then there exists a constant $\rho_1 > 0$ such that

$$\rho_1^{-1} \leq h_{K_i}/h_{K_j} \leq \rho_1.$$  

Moreover, if $c \subset \partial K$, there exist positive constants $c_1(\rho_1)$ and $c_2(\rho_1)$ independent of $h$ such that

$$c_1(\rho_1)h_K \leq h_c \leq c_2(\rho_1)h_K,$$  

where $h_c$ is the length of $c$.

For a positive integer $k$, we define the broken Sobolev space on $\mathcal{T}_h$,

$$H^k(\mathcal{T}_h) = \{v \in L^2(\Omega) : v|_K \in H^k(K), \quad \forall K \in \mathcal{T}_h\},$$  

equipped with the broken Sobolev norm and seminorm, respectively,

$$\|v\|_{k,h} = \left(\sum_{K \in \mathcal{T}_h} \|v\|_{k,K}^2\right)^{\frac{1}{2}}, \quad |v|_{k,h} = \left(\sum_{K \in \mathcal{T}_h} |v|_{k,K}^2\right)^{\frac{1}{2}}.$$  

For a given $\mathcal{T}_h$, we define the discontinuous finite element space $V_h$ by

$$V_h = \{v_h \in L^2(\Omega) : v_h|_K \in Z_r(K), \quad K \in \mathcal{T}_h\},$$  

where $Z_r(K)$ is the space $P_r(K)$ of polynomials of total degree $\leq r$, if $K$ is a triangle, or the space $Q_r(K)$ of polynomials of degree $\leq r$ in each variable, if $K$ is a quadrilateral, $1 \leq r$.

Next, we define the average and jump operators that are required for the DG method. To this end, we denote the set of interior edges of $\mathcal{T}_h$ by $\Gamma_I$ and the set of boundary edges by $\Gamma_B$. Furthermore, we define $\Gamma = \Gamma_I \cup \Gamma_B$. Let $K^+$ and $K^-$ be two adjacent elements of $\mathcal{T}_h$ which share a common edge $e$, $e = \partial K^+ \cap \partial K^-$. Furthermore, let $v$ and $q$ be scalar- and vector-valued functions, respectively, that are smooth inside each element $K^\pm$. $v^\pm$ and $q^\pm$ denote the traces of $v$ and $q$ on $e$ taken from within the interior of $K^\pm$, respectively. Then, the averages and jumps of $v$ and $q$ on $e$ are given by, respectively,

$$\{v\} = \frac{1}{2}(v^+ + v^-), \quad [v] = v^+_K + v^-_K;$$
Here onwards we don't specify the dependent of the functions $a$.

### 2.3. Discontinuous Galerkin method

The weak formulation of (2.1) is defined as

\[
\int \left[ a(x, x, u, q) \right] (v - u) \, dx + \int \left[ b(x, x, u, q) \right] \, dx = \int f(x) \, dx,
\]

where $a(x, x, u, q)$ and $b(x, x, u, q)$ are the bilinear and linear forms, respectively.

### Lemma 2.1

Let $\omega \in H^1(K)$ and $e$ an edge of $K \in T_h$. There exists a constant $C$ independent of $h_K$ such that

\[
\| \omega \|_{0,e}^2 \leq C(h^{-1}_e\| \omega \|_{0,K} + h_e\| \nabla \omega \|_{0,K}).
\]

From Lemma 2.1 and the inverse inequality [14], we have the following lemma.

### Lemma 2.2

Let $v_h \in Z_h(K)$. Then, there exists a positive constant $C$ independent of $h_K$ such that

\[
\| v_h \|_{0,e} \leq Ch^\frac{1}{2}_K \| v_h \|_{0,K}, \quad \| \nabla v_h \|_{0,e} \leq Ch^\frac{1}{2}_K \| v_h \|_{1,K}.
\]

The following integral form of the Taylor’s formula for $v \in \mathbb{R}$ and $p \in \mathbb{R}^2$ in terms of $u \in \mathbb{R}$ and $q \in \mathbb{R}^2$ will be used in the subsequent analysis:

\[
a_i(x, v, p) - a_i(x, u, q) = a_i(x, u, q)(v - u) + \tilde{a}_i(x, u, q)(p - q), \quad i = 0, 1, 2,
\]

where

\[
\tilde{a}_i(x, u, q) = \int_0^1 a_i(x, u, v^t, p^t) \, dt, \quad \tilde{a}_i(x, u, q) = \int_0^1 a_i(x, v^t, p^t) \, dt,
\]

\[
v^t = u + t(v - u), \quad p^t = q + t(p - q).
\]

### 2.3. Discontinuous Galerkin method

The weak formulation of (2.1) is defined as

\[
a(u, v) = (f, v), \quad \forall v \in H^1_0(\Omega),
\]

where

\[
a(u, v) = \int_\Omega (a(u, \nabla u) \cdot \nabla v + a_0(u, \nabla u) v) \, dx, \quad (f, v) = \int_\Omega f v \, dx.
\]

Here onwards we don’t specify the dependent of the functions $a$ and $a_0$ on $x$.

In order to present the incomplete interior penalty method, we introduce the form $a_h(\omega_h, v_h)$ for $\omega_h, v_h \in V_h$

\[
a_h(\omega_h, v_h) = \sum_{K \in T_h} \int_K (a(\omega_h, \nabla \omega_h) \cdot \nabla v_h + a_0(\omega_h, \nabla \omega_h) v_h) \, dx.
\]

(2.10)

\[
-\sum_{e \in E} \int_e \{ a(\omega_h, \nabla \omega_h) \} [v_h] \, ds + \sum_{e \in E} \int_e \{ a(\omega_h) \} [v_h] \, ds,
\]

where $\gamma$ is the discontinuity penalization parameter independent of $h_e$.

The incomplete interior penalty method for (2.1) is: find $u_h \in V_h$ such that

\[
a_h(u_h, v_h) = (f, v_h), \quad \forall v_h \in V_h.
\]

(2.11)

It will be shown that there is a parameter $\gamma_0 > 0$ independent of $h_e$ such that for $\gamma \geq \gamma_0$, the incomplete interior penalty method (2.11) possesses a unique solution.
We introduce the following so-called energy norms, see [30, 31] and [28], respectively

\[ ||v|| = \left( |v|^2 h + \sum_{e \in \Gamma} \gamma h_{e} \int_{e} |v|^2 ds \right)^{\frac{1}{2}}, \quad ||v||_- = \left( |v|^2 h_{1} + \sum_{e \in \Gamma} \frac{1}{h_{e}} \int_{e} |v|^2 ds \right)^{\frac{1}{2}}. \]

Note that the norm \( || \cdot || \) depends on the parameter \( \gamma \) and the norm \( || \cdot ||_- \) is independent of \( \gamma \).

We state the Poincaré-type inequalities on \( H^1(T_h) \). For a proof, we refer to [5].

**Lemma 2.3.** ([5]) Let \( v \in H^1(T_h) \). Then there exists a constant \( C > 0 \) independent of \( h \) and \( v \) such that

\[ ||v|| \leq C ||v||_- . \]

From Lemma 2.3, we know that for \( \gamma \geq 1 \)

\[ ||v|| \leq C ||v||_- . \] (2.12)

**3. Existence and uniqueness of DG solution**

In this section, using a result from the theory of monotone operators, we will show that the problem (2.11) has a unique solution. For this purpose, we first prove the form \( a_h(\cdot, \cdot) \) is bounded, Lipschitz-continuous and strongly monotone.

The following lemma gives the boundedness of the form \( a_h(\cdot, \cdot) \).

**Lemma 3.1.** There is a constant \( C > 0 \) such that for any \( \omega_h, v_h \in V_h \) and \( \gamma \geq 1 \)

\[ |a_h(\omega_h, v_h)| \leq C(1 + ||\omega_h||)||v_h||. \] (3.1)

**Proof.** From assumption (A), we know that there exists a constant \( c_0 \) such that

\[ |a_i(x, \xi)| \leq c_0 \left( 1 + \sum_{j=0}^{2} |\xi_j| \right), \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^3, \quad i = 0, 1, 2. \] (3.2)

From the definition (2.10) of the form \( a_h(\cdot, \cdot) \), we see that it suffices to bound each term on the left-hand side of (2.10) by the right-hand side of (3.1). In fact, by Cauchy-Schwarz inequality, (3.2) and (2.12), we have

\[ \left| \sum_{K \in T_h} \int_{K} (a(\omega_h, \nabla \omega_h) \cdot \nabla v_h + a_0(\omega_h, \nabla \omega_h) v_h) dx \right| \leq C(1 + ||\omega_h||)||v_h||. \] (3.3)
Similarly, we have

\[ \sum_{e \in \Gamma} \int_e \{a(\omega_h, \nabla \omega_h)\} |v_h| ds \]

\[ \leq \sum_{e \in \Gamma} \| \{a(\omega_h, \nabla \omega_h)\} \|_{0,e} \|v_h\|_{0,e} \]

\[ \leq \left( \sum_{e \in \Gamma} \frac{h_e}{\gamma} \| \{a(\omega_h, \nabla \omega_h)\} \|_{0,e}^2 \right) \left( \sum_{e \in \Gamma} \frac{\gamma}{h_e} \|v_h\|_{0,e}^2 \right)^{\frac{1}{2}} \]

(3.4)

\[ \leq C \left( \sum_{K \in \mathcal{T}_h} \sum_{e \in \partial K} h_e \left( h_e + \|\omega_h\|_{0,e}^2 + \|\nabla \omega_h\|_{0,e}^2 \right) \right)^{\frac{1}{2}} \|v_h\| \]

\[ \leq C(1 + \|\omega_h\| + \|\omega_{1,h}\|)\|v_h\| \]

\[ \leq C(1 + \|\omega_h\|)\|v_h\|. \]

Using Cauchy-Schwarz inequality, we get

\[ \left| \sum_{e \in \Gamma} \int_e [\omega_h][v_h] ds \right| \leq \left( \sum_{e \in \Gamma} \frac{\gamma}{h_e} \|\omega_h\|_{0,e}^2 \right) \left( \sum_{e \in \Gamma} \frac{\gamma}{h_e} \|v_h\|_{0,e}^2 \right)^{\frac{1}{2}} \]

(3.5)

\[ \leq \|\omega_h\| \|v_h\|. \]

Substituting (3.3), (3.4) and (3.5) into (2.10) completes the proof of this lemma. □

**Lemma 3.2.** The form \( a_h(\cdot, \cdot) \) is Lipschitz-continuous in its first argument for \( \gamma \geq 1 \)

(3.6)

\[ |a_h(\omega_1, v) - a_h(\omega_2, v) | \leq C ||\omega_1 - \omega_2|| \|v\|, \quad \forall \omega_1, \omega_2, v \in V_h. \]

**Proof.** It follows from the definition of \( a_h(\cdot, \cdot) \) that

\[ a_h(\omega_1, v) - a_h(\omega_2, v) = \sum_{K \in \mathcal{T}_h} \int_K (a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2)) \cdot \nabla v dx \]

\[ + \sum_{K \in \mathcal{T}_h} \int_K (a_0(\omega_1, \nabla \omega_1) - a_0(\omega_2, \nabla \omega_2)) \cdot \nabla v dx \]

\[ - \sum_{e \in \Gamma} \int_e \{a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2)\} [v] ds \]

\[ + \sum_{e \in \Gamma} \frac{\gamma}{h_e} \int_e [\omega_1 - \omega_2][v] ds \]

\[ = I_1 + \cdots + I_4. \]

Applying Cauchy-Schwarz inequality, (2.8), assumption (A) and (2.12), we deduce that

(3.8)

\[ |I_1| \leq C(||\omega_1 - \omega_2|| + |\omega_1 - \omega_{2,1,h}|v|_{1,h}) \leq C ||\omega_1 - \omega_2|| \|v\|. \]

Similarly, we have

(3.9)

\[ |I_2| \leq C ||\omega_1 - \omega_2|| \|v\|. \]
Using Cauchy-Schwarz inequality, (2.8), Lemma 2.2 and (2.12), we can bound \( I_3 \)
\[
|I_3| \leq \sum_{e \in T_h} \left( \frac{h}{\gamma} \int_e \{ a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2) \}^2 ds \right)^\frac{1}{2} \left( \frac{\gamma}{h_e} \int_e [v]^2 ds \right)^\frac{1}{2}
\]
\[
\leq \left( \sum_{e \in T_h} \frac{h}{\gamma} \int_e \{ a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2) \}^2 ds \right)^\frac{1}{2} \left( \sum_{e \in T_h} \frac{\gamma}{h_e} \int_e [v]^2 ds \right)^\frac{1}{2}
\]
\[
(3.10) \leq C \left( \sum_{K \in T_h} \frac{h_K}{\gamma} \int_{\partial K} \{ a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2) \}^2 ds \right)^\frac{1}{2} \||v||
\]
\[
\leq C \frac{1}{\gamma^2} \left( \sum_{K \in T_h} h_K (||\omega_1 - \omega_2||_{0, \partial K} + ||\nabla (\omega_1 - \omega_2)||_{0, \partial K}) \right)^\frac{1}{2} ||v||
\]
\[
\leq C \frac{1}{\gamma^2} (||\omega_1 - \omega_2|| + ||\omega_1 - \omega_2|_{1, h}) ||v||
\]
\[
\leq C' \frac{1}{\gamma^2} ||\omega_1 - \omega_2|| ||v||.
\]

The estimation of the fourth term \( I_4 \) on the right-hand side of (3.7) is easy
\[
|I_4| \leq \left( \sum_{e \in T_h} \frac{\gamma}{h_e} ||\omega_1 - \omega_2||_{0, e}^2 \right)^\frac{1}{2} \left( \sum_{e \in T_h} \frac{\gamma}{h_e} ||v||_{1, e}^2 \right)^\frac{1}{2}
\]
\[
(3.11) \leq C ||\omega_1 - \omega_2|| ||v||.
\]

Combining (3.8)-(3.11) with (3.7) yields the desired result (3.6). \(\square\)

**Lemma 3.3.** There exists a constant \( \gamma_0 > 1 \) such that for \( \gamma \geq \gamma_0 \), \( a_h(\cdot, \cdot) \) is strongly monotone in the sense that
\[
\frac{1}{2} \min(a, 1) ||\omega_1 - \omega_2||^2 \leq a_h(\omega_1, \omega_1 - \omega_2) - a_h(\omega_2, \omega_1 - \omega_2), \quad \forall \omega_1, \omega_2 \in V_h.
\]

**Proof.** Setting \( \omega = \omega_1 - \omega_2 \), from the definition of \( a_h(\cdot, \cdot) \), we have
\[
a_h(\omega_1, \omega) - a_h(\omega_2, \omega) = \sum_{K \in T_h} \int_K (a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2)) \cdot \nabla \omega \, dx
\]
\[
+ \sum_{K \in T_h} \int_K (a_0(\omega_1, \nabla \omega_1) - a_0(\omega_2, \nabla \omega_2)) \omega \, dx
\]
\[
- \sum_{e \in T_h} \int_e [a(\omega_1, \nabla \omega_1) - a(\omega_2, \nabla \omega_2)] [\omega] \, ds
\]
\[
+ \sum_{e \in T_h} \frac{\gamma}{h_e} \int_e [\omega]^2 \, ds
\]
\[
= J_1 + \cdots + J_4.
\]

In order to formulate \( a_i(\omega_1, \nabla \omega_1) - a_i(\omega_2, \nabla \omega_2), i = 0, 1, 2 \), we introduce the notation \( \zeta, \eta, \tau \in \mathbb{R}^3 \) and set
\[
g_i(t) = a_i(\zeta + t(\eta - \zeta)), \quad i = 0, 1, 2.
\]

Obviously, we have
\[
a_i(\zeta) = g_i(0), \quad a_i(\eta) = g_i(1), \quad g_i'(t) = \sum_{j=0}^2 \frac{\partial a_i}{\partial \zeta_j}(\zeta + t(\eta - \zeta))(\eta_j - \zeta_j).
\]
Theorem 3.5. Suppose $H$ is continuous on $L$ and strongly monotone, i.e., there exists $\rho > 0$ and $\rho > 0$ such that

$$J_{a_j}(\xi) = \frac{1}{2} \sum_{j=1}^{\infty} \int_{0}^{1} \frac{\partial a_j}{\partial \xi_j}(\xi + t(\eta - \zeta))(\eta_j - \zeta_j)dt.$$ 

Thus

$$\sum_{i=0}^{2} (a_i(\eta) - a_i(\zeta)) \tau_i = \sum_{i=0}^{2} \int_{0}^{1} \frac{\partial a_i}{\partial \xi_j}(\xi + t(\eta - \zeta))(\eta_j - \zeta_j)\tau_i dt.$$ 

Setting $\tau = \eta - \zeta$ in (3.14) and using the assumption (B), we immediately obtain

$$\alpha \sum_{i=1}^{2} (\eta_i - \zeta_i)^2 \leq \sum_{i=0}^{2} (a_i(\eta) - a_i(\zeta)) (\eta_i - \zeta_i).$$

Let $\eta_0 = \omega_1, \eta_1 = \frac{\partial}{\partial \omega_1}, \eta_2 = \frac{\partial}{\partial \omega_2}$ and $\zeta_0 = \omega_2, \zeta_1 = \frac{\partial}{\partial \omega_1}, \zeta_2 = \frac{\partial}{\partial \omega_2}$. Inserting them into (3.15), we can get the lower bound on $J_1 + J_2$ on the right-hand side of (3.12)

$$\alpha \sum_{K \in T_0} |\omega_1 - \omega_2|^2 \leq J_1 + J_2.$$ 

Then, from (3.16), the definitions of $||| \cdot |||$ and $J_4$ on the right-hand side of (3.12), we know that

$$\min(\alpha, 1) ||| \omega_1 - \omega_2 |||^2 \leq J_1 + J_2 + J_4.$$ 

From (3.10), we get the estimation of the third term on the right-hand side of (3.12)

$$|J_3| \leq C' ||| \omega_1 - \omega_2 |||^2.$$ 

Combining (3.17), (3.18) with (3.12) yields

$$\left( \min(\alpha, 1) - \frac{C'}{\gamma^2} \right) ||| \omega_1 - \omega_2 |||^2 \leq a_k(\omega_1, \omega) - a_k(\omega_2, \omega).$$

Then, choosing $\gamma_0$ such that $C'/\gamma_0^{1/2} \leq \frac{1}{2} \min(\alpha, 1)$, we obtain the desired result.

We conclude this section by proving the existence and uniqueness of the solution of (2.11) by means of a result from the theory of monotone operators. The proof is the same as that of Theorem 2.5 in [30]. However, for sake of completeness, we present the main steps.

We shall make use of the following result from the monotone operator theory (see Theorem 3.2.23 in [40]).

**Lemma 3.4.** ([40]) Let $H$ be a real Hilbert space with inner product $(\cdot, \cdot)_H$ and norm $|| \cdot ||_H$, and let $L$ be an operator from $H$ into itself. Suppose that $L$ is Lipschitz-continuous, i.e., there exists $\Lambda > 0$ such that

$$||L(w_1) - L(w_2)||_H \leq \Lambda ||w_1 - w_2||_H, \quad w_1, w_2 \in H,$$

and strongly monotone, i.e., there exists $\rho > 0$ such that

$$\rho ||w_1 - w_2||_H^2 \leq (L(w_1) - L(w_2), w_1 - w_2)_H.$$

Then, $L$ is a bijection of $H$ onto itself, and the inverse $L^{-1}$ of $L$ is Lipschitz-continuous on $H$:

$$||L^{-1}f - L^{-1}g||_H \leq (\Lambda/\rho)||f - g||_H, \quad \forall f, g \in H.$$

**Theorem 3.5.** Suppose $\gamma \geq \gamma_0$. Then the problem (2.11) has a unique solution.
Proof. Using the Riesz representation theorem from Hilbert space theory, we shall first rewrite (2.11) as an equivalent nonlinear operator equation \(\mathcal{L}(u_h) = f\) on \(\mathcal{H} \equiv V_h\). Then, applying Lemma 3.4, we deduce that this has a unique solution \(u_h\) in \(V_h\).

We know that \(V_h\) is a finite dimensional Hilbert space with the norm \(||| \cdot |||\) induced by the inner product \(\langle \cdot, \cdot \rangle\), where

\[
(3.20) \quad \langle \omega, v \rangle = \sum_{K \in T_h} \int_K \nabla \omega \cdot \nabla v \, dx + \sum_{e \in \Gamma} \frac{1}{h_e} \int_e [\omega] [v] \, ds.
\]

Note that \((f, v_h)\) is a bounded linear functional on \(V_h\). In fact, by Cauchy-Schwarz inequality and (2.12)

\[
|(f, v_h)| \leq ||f|| \, ||v_h|| \leq C_1 ||f|| \, ||v_h|| \leq C_2 ||v_h||, \quad C_2 = C_1 ||f||.
\]

Hence, by the Riesz representation theorem, there exists \(f \in V_h\) such that

\[
(f, v_h) = (f, v_h), \quad \forall v_h \in V_h.
\]

Given any \(\omega_h \in V_h\), we consider the following linear functional on \(V_h\)

\[
(3.21) \quad \Phi_{\omega_h} : \Phi_{\omega_h}(v_h) = a_h(\omega_h, v_h), \quad \forall v_h \in V_h.
\]

From Lemma 3.1, we see that the linear functional \(\Phi_{\omega_h}\) is bounded on \(V_h\), i.e.,

\[
|\Phi_{\omega_h}(v_h)| \leq C_3 (1 + ||\omega_h||) ||v_h|| \leq C_4 ||v_h||, \quad C_4 = C_3 (1 + ||\omega_h||).
\]

Then, by virtue of the Riesz representation theorem, there exists \(\mathcal{L}(\omega_h) \in V_h\) such that

\[
(3.22) \quad \Phi_{\omega_h}(v_h) = \langle \mathcal{L}(\omega_h), v_h \rangle, \quad \forall v_h \in V_h.
\]

As \(\omega_h\) passes through \(V_h\), (3.22) defines the mapping of \(V_h\) onto itself

\[
\omega_h \in V_h \mapsto \mathcal{L}(\omega_h) \in V_h.
\]

Thus, we rewrite (2.11) as a nonlinear equation: find \(u_h \in V_h\) such that \(\mathcal{L}(u_h) = f\).

From Lemma 3.2, we know that the mapping \(\omega_h \mapsto \mathcal{L}(\omega_h)\) is Lipschitz-continuous with respect to the norm \(||| \cdot |||\),

\[
|||\mathcal{L}(\omega_1) - \mathcal{L}(\omega_2)||| = \sup_{v_h \in V_h} \frac{|\langle \mathcal{L}(\omega_1) - \mathcal{L}(\omega_2), v_h \rangle|}{||v_h||}
= \sup_{v_h \in V_h} \frac{|\Phi_{\omega_1}(v_h) - \Phi_{\omega_2}(v_h)|}{||v_h||}
= \sup_{v_h \in V_h} \frac{|a_h(\omega_1, v_h) - a_h(\omega_2, v_h)|}{||v_h||}
\leq C_6 ||\omega_1 - \omega_2||, \quad \forall \omega_1, \omega_2 \in V_h.
\]

From (3.22), (3.21) and Lemma 3.3, we see that the mapping \(\mathcal{L}\) is strongly monotone

\[
\langle \mathcal{L}(\omega_1) - \mathcal{L}(\omega_2), \omega_1 - \omega_2 \rangle = \Phi_{\omega_1}(\omega_1 - \omega_2) - \Phi_{\omega_2}(\omega_1 - \omega_2)
= a_h(\omega_1, (\omega_1 - \omega_2)) - a_h(\omega_2, (\omega_1 - \omega_2))
\geq C_6 ||\omega_1 - \omega_2||^2.
\]

Then, from Lemma 3.4, we know that \(\mathcal{L}\) is a bijection of \(V_h\) onto itself. Hence, for any \(f \in V_h\), the equation \(\mathcal{L}(u_h) = f\), which is equivalent to (2.11), has a unique solution \(u_h \in V_h\).
4. A posteriori error estimates

In this section, we present the a posteriori error estimates on the error in the energy norm $\| \cdot \|_\cdot$ for the DG method.

4.1. A reliable a posteriori error bound. In this subsection, we propose the residual-based a posteriori error estimators and derive the computable upper bounds on the error $(u - u_h)$ in the broken $H^1$-seminorm and in the energy norm $\| \cdot \|_\cdot$.

In our a posteriori error analysis, an important technique, which was employed in [31, 33, 35, 36, 37], is the decomposition of the discontinuous finite element space $V_h$ into two orthogonal subspaces: a conforming part $V_h^c = \cap_{x} H^1_0(\Omega)$, and a nonconforming part $V_h^\perp$ defined as the orthogonal complement of $V_h^c$ in $V_h$ with respect to the energy norm $\| \cdot \|_\cdot$, i.e.,

$$V_h = V_h^c \oplus V_h^\perp.$$

Based on this setting, the discontinuous Galerkin approximation $u_h$ may be split accordingly,

$$u_h = u_h^c + u_h^\perp. \tag{4.1}$$

The following lemma describes the approximations of discontinuous finite element functions by conforming ones, which have been established in Theorem 2.2 and Theorem 2.3 in [35] for conforming and nonconforming meshes (see also Theorem 2.1 in [36]).

**Lemma 4.1.** For each $v_h \in V_h$, there is a constant $C^0$ independent of $h$ and $v_h$ such that

$$|v_h|_{1,h}^2 = |v_h - v_h^c|^2_{0,h} \leq C^0 \sum_{e} h_e^{-1} \|v_h\|_{0,e}^2,$$

and

$$\|v_h - v_h^c\|^2 \leq C^0 \sum_{e} h_e \|v_h\|_{0,e}^2. \tag{4.2}$$

From Lemma 4.1, we know that

$$\sum_{K} h_K^{-2} \|v_h - v_h^c\|_{0,K}^2 + |v_h^c|_{1,h}^2 \leq C \||v_h\|_\cdot, \quad \forall v_h \in V_h. \tag{4.4}$$

Houston, Süli and Wihler [33] obtained an approximation result in the $hp$-discontinuous finite element space on the nonconforming mesh. Here we state it in the case of $h$-discontinuous finite element space.

**Lemma 4.2.** For each $v \in H^1_0(\Omega)$, there exists a function $v_I \in V_h$ such that

$$\sum_{K} (h_K^{-2} \|v - v_I\|^2_{0,K} + |v - v_I|_{1,K}^2 + h_K^{-1} \|v - v_I\|^2_{0,\partial K}) \leq C |v|^2_{1}.\tag{4.3}$$

The following approximation result will also be used in the a posteriori error analysis.

**Lemma 4.3.** For each $v \in H^1_0(\Omega)$, there exists a constant $C$ independent of $h$ such that

$$\sum_{K} (h_K^{-2} \|v - v_I\|^2_{0,K} + |v - v_I|_{1,K}^2 + h_K^{-1} \|v - v_I\|^2_{0,\partial K}) \leq C |v|^2_{1},$$

$$\sum_{K} h_K^{-2} \|v - v_I\|^2_{0,K} + |v - v_I|_{1,K}^2 + h_K^{-1} \|v - v_I\|^2_{0,\partial K} \leq C |v|^2_{1}.$$
where \( v_f^c \in V_h^c \) is the conforming part of \( v_f \) and \( v_I \) is given in Lemma 4.2.

**Proof.** Using Lemma 4.1, \([v] = 0\) and Lemma 4.2, we get

\[
\sum_{K \in T_h} (h_K^{-2}|v_I - v_f^c|^2_{0,K} + |v_I - v_f^c|^2_{1,K}) \leq C \sum_{e \in F} h_e^{-1}|[v_I]|^2_{0,e} \\
= C \sum_{e \in F} h_e^{-1}|[v - v_I]|^2_{0,e} \\
\leq C \sum_{K \in T_h} h_K^{-2}|v - v_I|^2_{0,\partial K} \\
\leq C|v_f|^2_{1_I}.
\]

It follows from Lemma 2.2 and (4.5) that

\[
\sum_{K \in T_h} h_K^{-2}|v_I - v_f^c|^2_{0,\partial K} \leq C \sum_{K \in T_h} h_K^{-2}|v_I - v_f^c|^2_{0,K} \leq C|v_f|^2_{1_I}.
\]

Then the proof is completed by combining (4.5) with (4.6).

The following lemma gives a representation of the error \( u - u_h \), which plays a key role in the a posteriori error estimation. A novel contribution in Lemma 4.4 is that, by introducing the functions \( v = u - u_h^c \) and \( v_f^c \) which satisfy \([v] = [v_f] = 0\), the penalty parameter \( \gamma \) disappears from the representation of the error. This will lead to the upper bound without \( \gamma \) for the error in the broken \( H^1 \)-seminorm.

**Lemma 4.4.** Assume that \( u \in H^1(\Omega) \) and \( u_h \in V_h \) are the solutions of (2.1) and (2.11), respectively. Then for the error \( u - u_h \) and \( v = u - u_h^c \in H_0^1(\Omega) \), we have

\[
\begin{align*}
\mathbb{I} &= \sum_{K \in T_h} \int_K (a(u, \nabla u) - a(u_h, \nabla u_h)) \cdot \nabla (u - u_h) \, dx \\
&\quad + \sum_{K \in T_h} \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h)) (u - u_h) \, dx \\
&= \sum_{K \in T_h} \int_K (f + \nabla : a(u_h, \nabla u_h) - a(u, \nabla u_h)) (v - v_f^c) \, dx \\
&\quad - \sum_{e \in F} \int_e [a(u_h, \nabla u_h)](v - v_f^c) \, ds \\
&\quad - \sum_{K \in T_h} \int_K (a(u, \nabla u) - a(u_h, \nabla u_h)) \cdot \nabla u_h^i \, dx \\
&\quad - \sum_{K \in T_h} \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h)) u_h^i \, dx \\
&= K_1 + K_2 + K_3 + K_4.
\end{align*}
\]

where \( v_f^c \in V_h^c \) is the conforming part of \( v_f \) and \( v_I \) is given in Lemma 4.2.

**Proof.** For \( v = u - u_h^c \in H_0^1(\Omega) \), we have \( v_f \in V_h \) and \( v_f^c \in V_h^c \). Then, from (2.9) and (2.11), we get

\[
\begin{align*}
a(u, v) - a_h(u_h, v) &= (f, v) - a_h(u_h, v) = (f, v - v_f^c) - a_h(u_h, v - v_f^c).
\end{align*}
\]
Since \( v - v^e_f \in H^1_0(\Omega) \), we have \( [v - v^e_f] = 0 \). Then, it follows from the definition of \( a_h(\cdot, \cdot) \) and Green’s formula that

\[
a_h(u_h, v - v^e_f) = \sum_{K \in T_h} \int_K \mathbf{a}(u_h, \nabla u_h) \cdot \nabla (v - v^e_f) \, dx + \sum_{K \in T_h} a_0(u_h, \nabla u_h)(v - v^e_f) \, dx
\]

\[
= - \sum_{K \in T_h} \int_K \nabla \cdot \mathbf{a}(u_h, \nabla u_h)(v - v^e_f) \, dx + \sum_{K \in T_h} \int_{\partial K} \mathbf{a}(u_h, \nabla u_h) \cdot \nu_K (v - v^e_f) \, ds
\]

\[
+ \sum_{K \in T_h} \int_K a_0(u_h, \nabla u_h)(v - v^e_f) \, dx.
\]

Using the following identity,

\[
\sum_{K \in T_h} \int_{\partial K} \mathbf{a}(u_h, \nabla u_h) \cdot \nu_K (v - v^e_f) \, ds = \sum_{e \in \Gamma} \int_e \{ \mathbf{a}(u_h, \nabla u_h) \} [v - v^e_f] \, ds
\]

\[
+ \sum_{e \in \Gamma} \int_e \{ \mathbf{a}(u_h, \nabla u_h) \} \{ v - v^e_f \} \, ds,
\]

(4.9)

and the fact \([v - v^e_f] = 0\), we have

\[
\sum_{K \in T_h} \int_{\partial K} \mathbf{a}(u_h, \nabla u_h) \cdot \nu_K (v - v^e_f) \, ds = \sum_{e \in \Gamma} \int_e \{ \mathbf{a}(u_h, \nabla u_h) \} (v - v^e_f) \, ds.
\]

Then

\[
a_h(u_h, v - v^e_f) = - \sum_{K \in T_h} \int_K \nabla \cdot \mathbf{a}(u_h, \nabla u_h)(v - v^e_f) \, dx
\]

\[
+ \sum_{e \in \Gamma} \int_e \{ \mathbf{a}(u_h, \nabla u_h) \} (v - v^e_f) \, ds
\]

\[
+ \sum_{K \in T_h} \int_K a_0(u_h, \nabla u_h)(v - v^e_f) \, dx.
\]

Inserting (4.10) into (4.8) gives

\[
a(u, v) - a_h(u_h, v) = \sum_{K \in T_h} \int_K (f + \nabla \cdot \mathbf{a}(u_h, \nabla u_h) - a_0(u_h, \nabla u_h))(v - v^e_f) \, dx
\]

\[
- \sum_{e \in \Gamma} \int_e \{ \mathbf{a}(u_h, \nabla u_h) \} (v - v^e_f) \, ds.
\]

(4.11)

On the other hand, by the decomposition \( u_h = u_h^0 + u_h^e \) and \( v = u - u_h^e \), we know that \( v = (u - u_h^0) + u_h^e \). Then, it follows from the definition of \( a_h(\cdot, \cdot) \) and \([v] = 0\)
that

\[ a(u, v) - a_h(u_h, v) = \sum_{K \in T_h} \int_K (a(u, \nabla u) - a(u_h, \nabla u_h)) \cdot \nabla v \text{d}x + \sum_{K \in T_h} \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h)) v \text{d}x \]

\[ = \sum_{K \in T_h} \int_K (a(u, \nabla u) - a(u_h, \nabla u_h)) \cdot \nabla (u - u_h) \text{d}x + \sum_{K \in T_h} \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h)) (u - u_h) \text{d}x \]

\[ = (4.12) \]

Combining (4.11) with (4.12) completes the proof.

Motivated by the above lemma, we introduce the following locally computable quantities which will be used in the definition of the residual-based a posteriori error estimators.

**Definition 4.1.** On each element \( K \in T_h \) and \( e \in \Gamma \), define the element residual and the edge residuals by, respectively,

\[ R_K = f + \nabla \cdot a(u_h, \nabla u_h) - a_0(u_h, \nabla u_h), \quad J_{e,1} = [a(u_h, \nabla u_h)]_e, \quad J_{e,2} = [u_h]_e \]

and define the local error estimators \( \eta^2_K, \eta^2_{e,1} \) and \( \eta^2_{e,2} \) by

\[ \eta^2_K = h^2_K ||R_K||_{0,K}, \quad \eta^2_{e,1} = h_e ||J_{e,1}||_{0,e}, \quad \eta^2_{e,2} = h_e^{-1} ||J_{e,2}||_{0,e}^2 \]

Define the global error estimators by

\[ \eta_R = \left( \sum_{K \in T_h} \eta^2_K \right)^{\frac{1}{2}}, \quad \eta_{I,1} = \left( \sum_{e \in \Gamma} \eta^2_{e,1} \right)^{\frac{1}{2}}, \quad \eta_{I,2} = \left( \sum_{e \in \Gamma} \eta^2_{e,2} \right)^{\frac{1}{2}}. \]

**Remark 4.1.** Our estimator \( \eta_{I,2} \) is independent of the parameter \( \gamma \).

We are now in position to develop a reliable estimate for the error \( |u - u_h| \) in the broken \( H^1 \)-seminorm for the DG method.

**Theorem 4.5.** Assume that \( u \) and \( u_h \) are the solutions of (2.1) and (2.11), respectively. Then

\[ |u - u_h|_{1,h} \leq C (\eta_R + \eta_{I,1} + \eta_{I,2}) \]

**Proof.** Similar to the derivation of (3.16), we can get the lower bound on the left-hand side of (4.7)

\[ \alpha |u - u_h|^2_{1,h} = \alpha \sum_{K \in T_h} |u - u_h|^2_{1,K} \leq \Pi. \]

Next, we will estimate the terms on the right-hand side of (4.7) separately. First, since \( v = u - u_h^e \) and \( u_h = u_h^e + u_h^i \), we have by the triangle inequality

\[ |v|_1 = |(u - u_h^e - u_h^i)| + |u_h^i| \leq |u - u_h|_{1,h} + |u_h^i|_{1,h}. \]
Applying Cauchy-Schwarz inequality, Lemma 4.3 and (4.15) lead to

\[ |K_1| \leq \left( \sum_{K \in T_h} h_K^2 |R_K|^2 ||u||^2_{0,K} \right)^{\frac{1}{2}} \left( \sum_{K \in T_h} h_K^2 |v - v_f|^2 ||u||^2_{0,K} \right)^{\frac{1}{2}} \]

(4.16)

\[ \leq C'_1 \eta_R |v|_1 \]

\[ \leq C'_1 \eta_R \left( |u - u_h|_{1,h} + |u_h^k|_{1,h} \right). \]

Then, using the following generalized arithmetic-geometric inequality

\[ c a b \leq \frac{c}{2} a^2 + \frac{c}{2} b^2, \quad \forall c > 0, \]

and the arithmetic-geometric mean inequality to estimate the terms including \( |u - u_h|_{1,h} \) and \( |u_h^k|_{1,h} \) on the right-hand side of (4.16), respectively, we have

\[ |K_1| \leq C'_1 \eta_{J,1} |v|_1 \]

\[ \leq C'_2 \eta_{J,1} (|u - u_h|_{1,h} + |u_h^k|_{1,h}). \]

(4.18)

Using Cauchy-Schwarz inequality, Lemma 4.3, (2.6), (4.15), (4.17) and the arithmetic-geometric mean inequality gives

\[ |K_2| \leq \sum_{e \in \Gamma_h} ||a(u_h, \nabla u_h)||_{0,e} \cdot ||v - v_f||_{0,e} \]

\[ \leq C \left( \sum_{e \in \Gamma_h} h_e ||a(u_h, \nabla u_h)||_{0,e} \right)^{\frac{1}{2}} \left( \sum_{e \in \Gamma_h} \frac{1}{h_e} ||v - v_f||^2_{e,0,e} \right)^{\frac{1}{2}} \]

(4.19)

\[ \leq C'_2 \eta_{J,1} |v|_1 \]

\[ \leq C'_2 \eta_{J,1} (|u - u_h|_{1,h} + |u_h^k|_{1,h}). \]

Applying Cauchy-Schwarz inequality, (2.8) and assumption (A), we have

\[ |K_3| + |K_4| \leq \sum_{K \in T_h} ||u - u_h||_{1,K} ||u_h^k||_{1,K} \leq C ||u - u_h||_{1,h} ||u_h^k||_{1,h}. \]

Noting that \( |v| = 0 \), by virtue of Lemma 2.3 we have that

\[ ||u - u_h||_{1,h} \leq |u - u_h|_{1,h} + ||u - u_h|| \]

\[ \leq |u - u_h|_{1,h} + C ||u - u_h||. \]

(4.21)

Since \( |u_h^k| = 0 \), we have \( |u_h^k| = |u_h^+ + u_h^-| = |u_h| \). Then, by Lemmas 2.3 and 4.1

\[ ||u_h^+||_{1,h} \leq |u_h^+|_{1,h} + |u_h^-|_{1,h} \leq |u_h^+|_{1,h} + ||u_h^-|| \]

\[ \leq C |u_h^+|_{1,h} + C \left( \sum_{h,e} \frac{1}{h_e} ||u_h^-||^2_{0,e} \right)^{\frac{1}{2}} \]

(4.22)
Substituting (4.21) and (4.22) into (4.20) and using (4.17) gives

\[ |K_3| + |K_4| \leq C_3' |u - u_h|_{1,h} \eta_{J,2} + C_3'' \eta_{J,2}^2 \]

(4.23)

\[ \leq \frac{\epsilon_3}{2} |u - u_h|_{1,h}^2 + \frac{C_3''}{2} \eta_{J,2}^2 + C_3' \eta_{J,2}^2. \]

From (4.14), (4.7), (4.18), (4.19), (4.23) and Lemma 4.1, we have

\[ \alpha |u - u_h|_{1,h}^2 \leq \frac{\epsilon_1 + \epsilon_2 + \epsilon_3}{2} |u - u_h|_{1,h}^2 + \left( \frac{C_1'^2}{2 \epsilon_1} + \frac{C_1'}{2} \right) \eta_R^2 \]

\[ + \left( \frac{C_2'^2}{2 \epsilon_2} + \frac{C_2'}{2} \right) \eta_{J,1}^2 + \left( \frac{C_1'^2}{2 \epsilon_1} + \frac{C_2'^2}{2 \epsilon_3} + C_3' \right) \eta_{J,2}^2. \]

(4.24)

Taking \( \epsilon_i \) sufficiently small such that

\[ \frac{\epsilon_1 + \epsilon_2 + \epsilon_3}{2} \leq \frac{\alpha}{2}. \]

(4.25)

Then, combining (4.25) with (4.24) yields

\[ \frac{\alpha}{2} |u - u_h|_{1,h}^2 \leq \left( \frac{C_1'^2}{2 \epsilon_1} + \frac{C_1'}{2} \right) \eta_R^2 + \left( \frac{C_2'^2}{2 \epsilon_2} + \frac{C_2'}{2} \right) \eta_{J,1}^2 \]

\[ + \left( \frac{C_1'^2}{2 \epsilon_1} + \frac{C_2'^2}{2 \epsilon_3} + C_3' \right) \eta_{J,2}^2, \]

which completes the proof. \( \square \)

Using Theorem 4.5, we immediately obtain the upper bound on the error in the energy norm \( ||| \cdot |||_\varepsilon \).

**Theorem 4.6.** Assume that \( u \) and \( u_h \) are the solutions of (2.1) and (2.11), respectively. Then

\[ |||u - u_h|||_\varepsilon \leq C (\eta_R + \eta_{J,1} + \eta_{J,2}). \]

**Proof.** It follows from \( |u| = 0 \) that

\[ \sum_{e \in \Gamma} h_e^{-1} |||u - u_h|||_{0,e}^2 = \sum_{e \in \Gamma} h_e^{-1} |||u_h|||_{0,e}^2 = \eta_{J,2}^2. \]

(4.26)

Then, the desired result (4.26) follows from the definition of \( ||| \cdot |||_\varepsilon \), Theorem 4.5 and (4.27). \( \square \)

**4.2. Efficiency.** In this subsection, we derive the local lower bounds on the error \((u - u_h)\).

To derive the bounds, we introduce the oscillations of the element residual \( R_K \) and the edge residual \( J_{e,1} \) as

\[ \text{osc}_{R,K}^2(u_h) = h_K^2 |R_K - \Pi_K R_K|_{0,K}^2, \quad \text{osc}_{J,e}^2(u_h) = h_e |J_{e,1} - \Pi_e J_{e,1}|_{0,e}^2, \]

where \( \Pi_K R_K \) is the element-wise \( L^2 \)-projection of \( R_K \) onto the space \( Z_{r-1}(K) \) and \( \Pi_e J_{e,1} \) is the \( L^2 \)-projection of \( J_{e,1} \) onto the space \( P_{r-1}(e) \). We denote the total oscillation by

\[ \text{osc}_h(u_h) = \left( \sum_{K \in \mathcal{T}_h} \text{osc}_{R,K}^2(u_h) + \sum_{e \in \Gamma} \text{osc}_{J,e}^2(u_h) \right)^{\frac{1}{2}}. \]

As auxiliary tools, we need the following bubble functions ([54]). For each triangle \( K \in \mathcal{T}_h \), denote by \( \lambda_{K,1}, \lambda_{K,2}, \lambda_{K,3} \) the barycentric co-ordinates. Define the
triangle-bubble function $b_K$ by

\begin{equation}
(4.28)
b_K = \begin{cases}
27\lambda_K,1\lambda_K,2\lambda_K,3, & \text{on } K, \\
0, & \text{on } \Omega\setminus K.
\end{cases}
\end{equation}

Given an interior edge $e = \partial K_1 \cap \partial K_2$ and $\omega_e = K_1 \cup K_2$, enumerate the vertices of $K_1$ and $K_2$ such that the vertices of $e$ are numbered first. We then define the edge-bubble function $b_e$ by

\begin{equation}
(4.29)
b_e = \begin{cases}
4\lambda_K,1\lambda_K,2, & \text{on } K_i, i = 1, 2, \\
0, & \text{on } \Omega\setminus \omega_e.
\end{cases}
\end{equation}

It is easy to see that $\text{supp } b_K \subset K, 0 \leq b_K \leq 1, \text{ supp } b_e \subset \omega_e, 0 \leq b_e \leq 1$.

Now, we give the lower bounds for the error indicators $\eta_K, \eta_{e,1}$ and $\eta_{e,2}$.

**Theorem 4.7.** Assume that $u \in H^1(\Omega)$ and $u_h \in V_h$ are the solutions of (2.1) and (2.11), respectively. Then, we have the following local lower bounds on the error $u - u_h$:

(i) for each element $K \in T_h$,

\begin{equation}
(4.30)
\eta_K \leq C \left( ||u - u_h||_{1,K} + \text{osc}_{R,K}(u_h) \right).
\end{equation}

(ii) for $e = \partial K_i \cap \partial K_j$ and $\omega_e = K_1 \cup K_2$,

\begin{equation}
(4.31)
\eta_{e,1} \leq C \left( ||u - u_h||_{1,h,\omega_e} + (\text{osc}_{R,K_i}(u_h) + \text{osc}_{R,K_j}(u_h)) + \text{osc}_{J,e}(u_h) \right).
\end{equation}

(iii) for $e \in \Gamma$,

\begin{equation}
(4.32)
\eta_{e,2} = h_e^{-\frac{1}{2}} ||u - u_h||_{0,e}.
\end{equation}

**Proof.** We present proof of the three assertions separately. Assertion (i): By triangle inequality,

\begin{equation}
(4.33)
h_K|| R_K||_{0,K} \leq h_K|| \Pi_K R_K||_{0,K} + h_K|| R_K - \Pi_K R_K||_{0,K}
\end{equation}

Thus, we only estimate $h_K|| \Pi_K R_K||_{0,K}$ in the following. Since $b_K > 0$ on int($K$), $(\int_K (b_K)^2 dx)^{1/2}$ defines a norm on $L^2(K)$, equivalent to the $L^2$ norm on $P_k(K)$ for any fixed $k$. Thus, there exists a constant $c'_1 > 0$ independent of $h_K$ such that

\begin{equation}
(4.34)
c'_1|| \Pi_K R_K||_{0,K}^2 \leq \int_K (\Pi_K R_K)^2 b_K dx.
\end{equation}
From the definition of $R_K$, (2.1), Green’s formula and $\text{supp}(b_K \Pi_K R_K) \subset K$, we get

\[
\int_K (\Pi_K R_K)^2 b_K dx = \int_K R_K (b_K \Pi_K R_K) dx + \int_K (\Pi_K R_K - R_K) (b_K \Pi_K R_K) dx
\]

\[
= \int_K (f + \nabla \cdot a(u_h, \nabla u_h) - a_0(u_h, \nabla u_h))(b_K \Pi_K R_K) dx
\]

\[
+ \int_K (\Pi_K R_K - R_K)(b_K \Pi_K R_K) dx
\]

\[
= \int_K (-\nabla \cdot a(u, \nabla u) + \nabla \cdot a(u_h, \nabla u_h))(b_K \Pi_K R_K) dx
\]

\[
(4.34)
\]

\[
+ \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h))(b_K \Pi_K R_K) dx
\]

\[
+ \int_K (\Pi_K R_K - R_K)(b_K \Pi_K R_K) dx
\]

\[
= \int_K (a(u, \nabla u) - a(u_h, \nabla u_h)) \cdot \nabla (b_K \Pi_K R_K) dx
\]

\[
+ \int_K (a_0(u, \nabla u) - a_0(u_h, \nabla u_h))(b_K \Pi_K R_K) dx
\]

\[
+ \int_K (\Pi_K R_K - R_K)(b_K \Pi_K R_K) dx
\]

\[
= Q_1 + Q_2 + Q_3.
\]

Using (2.8), Cauchy-Schwarz inequality, assumption (A) and the inverse inequality [14], we get

\[
|Q_1| + |Q_2| \leq C ||u - u_h||_1,K ||b_K \Pi_K R_K||_{1,K}
\]

\[
(4.35)
\]

\[
\leq Ch_K^{-1} ||u - u_h||_1,K ||b_K \Pi_K R_K||_{0,K}
\]

\[
\leq Ch_K^{-1} ||u - u_h||_1,K ||\Pi_K R_K||_{0,K}.
\]

Using Cauchy-Schwarz inequality and $\max_{x \in K} b_K(x) = 1$, we have

\[
|Q_3| \leq ||R_K - \Pi_K R_K||_{0,K} ||b_K \Pi_K R_K||_{0,K}
\]

\[
(4.36)
\]

\[
\leq ||R_K - \Pi_K R_K||_{0,K} ||\Pi_K R_K||_{0,K}.
\]

Combining (4.34), (4.35), (4.36) with (4.33) yields

\[
h_K ||\Pi_K R_K||_{0,K} \leq C ||u - u_h||_{1,K} + Ch_K ||R_K - \Pi_K R_K||_{0,K}.
\]

The desired result follows from (4.32) and the above inequality.

Assertion (ii): Let $e = \partial K_1 \cap \partial K_2$ and suppose that $e$ is a full edge of both $K_1$ and $K_2$. If $e$ is not a full edge of one of the triangles, we can prove this assertion as in [35]. By triangle inequality once more,

\[
(4.37)
\]

\[
h_e^2 ||J_{e,1}||_{0,e} \leq h_e^2 ||\Pi_e J_{e,1}||_{0,e} + h_e^2 ||J_{e,1} - \Pi_e J_{e,1}||_{0,e}.
\]

Similar to (4.33), we have

\[
(4.38)
\]

\[
c_e^2 ||\Pi_e J_{e,1}||_{0,e} \leq \int_e (\Pi_e J_{e,1})^2 b_e ds = R.
\]

Extend $\Pi_e J_{e,1}$ to a function $\varphi$ defined over $\omega_e$ by extending by constants along lines normal to $e$. From the definition of $b_e$, we know that $b_e \varphi \in H_0^1(\omega_e)$. Then, by
Green’s formula and (2.1), we rewrite the term on the right-hand side of (4.38) as follows

\[ R = \int_{\omega_e} (J_{e,1})(b_e \Pi_{c,1}) \text{d}s + \int_{\omega_e} (\Pi_{c,1} - J_{e,1})(b_e \Pi_{c,1}) \text{d}s \]

\[ = \int_{\omega_e} a(u_h, \nabla u_h) \cdot \nabla (b_e \varphi) \text{d}x + \int_{\omega_e} \nabla \cdot a(u_h, \nabla u_h)(b_e \varphi) \text{d}x \]

\[ + \int_{\omega_e} (\Pi_{c,1} - J_{e,1})(b_e \Pi_{c,1}) \text{d}s \]

\[ = \int_{\omega_e} (a(u_h, \nabla u_h) - a(u, \nabla u)) \cdot \nabla (b_e \varphi) \text{d}x + \int_{\omega_e} a(u, \nabla u) \cdot \nabla (b_e \varphi) \text{d}x \]

\[ + \int_{\omega_e} \nabla \cdot a(u_h, \nabla u_h)(b_e \varphi) \text{d}x + \int_{\omega_e} (\Pi_{c,1} - J_{e,1})(b_e \Pi_{c,1}) \text{d}s \]

\[ = \int_{\omega_e} (a(u_h, \nabla u_h) - a(u, \nabla u)) \cdot \nabla (b_e \varphi) \text{d}x + \int_{\omega_e} (a_0(u_h, \nabla u_h) - a_0(u, \nabla u))(b_e \varphi) \text{d}x \]

\[ + \int_{\omega_e} (\Pi_{c,1} - J_{e,1})(b_e \Pi_{c,1}) \text{d}s \]

\[ = R_1 + R_2 + R_3 + R_4, \]

where \( \nabla_h \varphi \) is the function whose restriction to element \( K \in T_h \) is equal to \( \nabla \varphi \).

From the definitions of \( b_e \) and \( \varphi \), we know that

(4.39) \[ ||b_e \varphi||_{\omega_e}^2 \leq ||\varphi||_{\omega_e}^2 = \int_{\omega_e} (\Pi_{c,1})^2 l(s) \text{d}s \leq h||\Pi_{c,1}||_{0,e}^2, \]

where \( l(s) \) is the length of line segment which is perpendicular to the edge \( e \) and intersects the boundary of \( \omega_e \). Using (2.8), Cauchy-Schwarz inequality, inverse inequality [14] and (4.39), we have

(4.40) \[ |R_1| + |R_2| \leq C ||u - u_h||_{1,h,\omega_e} ||b_e \varphi||_{1,\omega_e} \]

\[ \leq Ch^{-1} ||u - u_h||_{1,h,\omega_e} ||b_e \varphi||_{0,\omega_e} \]

\[ \leq Ch^{-1/2} ||u - u_h||_{1,h,\omega_e} ||\Pi_{c,1}||_{0,e}. \]

By Cauchy-Schwarz inequality and (4.39)

(4.41) \[ |R_3| \leq ||f + \nabla h \cdot a(u_h, \nabla u_h) - a_0(u_h, \nabla u_h)||_{0,\omega_e} ||b_e \varphi||_{0,\omega_e} \]

\[ \leq Ch^{-1} ||f + \nabla h \cdot a(u_h, \nabla u_h) - a_0(u_h, \nabla u_h)||_{0,\omega_e} ||\Pi_{c,1}||_{0,e}. \]

Applying Cauchy-Schwarz inequality gives

(4.42) \[ |R_4| \leq ||\Pi_{c,1} - J_{e,1}||_{0,e} ||b_e \Pi_{c,1}||_{0,e} \leq ||\Pi_{c,1} - J_{e,1}||_{0,e} ||\Pi_{c,1}||_{0,e}. \]

Combining (4.38), the equality \( R = R_1 + R_2 + R_3 + R_4 \), (4.40)-(4.42), we have

(4.43) \[ ||\Pi_{c,1}||_{0,e} \leq Ch^{-1/2} ||u - u_h||_{1,h,\omega_e} + Ch^{-1/2} ||f + \nabla h \cdot a(u_h, \nabla u_h) - a_0(u_h, \nabla u_h)||_{0,\omega_e} \]

\[ + ||\Pi_{c,1} - J_{e,1}||_{0,e}. \]
Obviously, multiplying $h^\frac{1}{2}$ on both sides of (4.43), and applying (4.37) and assertion (i) yield the result.

Assertion (iii): This is a simple consequence of the fact that $|u| = 0$, $\forall c \in \Gamma$. □

Using Theorems 4.7 and Lemma 2.3, we have

**Theorem 4.8.** Assume that $u \in H^1(\Omega)$ and $u_h \in V_h$ are the solutions of (2.1) and (2.11), respectively. Then, we have the following a posteriori lower bounds on the error $u - u_h$ in the energy norms $||| \cdot |||$.

$$\eta_R + \eta_{J,1} + \eta_{J,2} \leq C|||u - u_h||| - + \text{Cosc}_h(u_h).$$

**Remark 4.2.** Since osc$_h(u_h)$ is a higher order term, from Theorems 4.6 and 4.8, we see that $\eta_R + \eta_{J,1} + \eta_{J,2}$ is a reliable and efficient a posteriori error estimator of $|||u - u_h||| -$.

**5. A priori error estimate**

In this section, we derive the error estimate in the energy norm $|||u - u_h|||$ of (2.11). We first prove an abstract lemma.

**Lemma 5.1.** Assume that $u \in H^1(\Omega)$ and $u_h \in V_h$ are the solutions of (2.1) and (2.11), respectively. Then, for $\gamma \geq \gamma_0$, there exists a positive constant $C$ independent of $h$ and $\gamma$ such that

$$|||u - u_h||| \leq C \inf_{v_h \in V_h} \left( |||u - v_h||| + \sup_{\varphi \in V_h \setminus \{0\}} \frac{(f, \varphi - \varphi_h^c \gamma) - a_h(v_h, \varphi - \varphi_h^c)}{|||\varphi|||} \right),$$

where $\varphi_h^c$ is the conforming part of $\varphi_h$.

**Proof.** Choose $v_h \in V_h$ such that $v_h \neq u_h$. Let $\phi_h = u_h - v_h$. From Lemma 3.3, (2.11) and (2.9), we have

$$C|||u_h - v_h|||^2 \leq a_h(u_h, \phi_h) - a_h(v_h, \phi_h)$$

$$= (f, \phi_h) - a_h(v_h, \phi_h)$$

$$= a(u, \phi_h^c) - a_h(v_h, \phi_h^c) + (f, \phi_h - \phi_h^c) - a_h(v_h, \varphi_h - \phi_h^c),$$

where $\phi_h^c \in V_h^c$ is the conforming part of $\phi_h$. Therefore,

$$|||u_h - v_h||| \leq C \left( a(u, \phi_h^c) - a_h(v_h, \phi_h^c) \right) + \frac{(f, \phi_h - \phi_h^c) - a_h(v_h, \varphi_h - \phi_h^c)}{|||u_h - v_h|||}.$$

Since $\phi_h^c \in V_h \cap H^1_0(\Omega)$, by the definitions of $a(\cdot, \cdot)$ and $a_h(\cdot, \cdot)$, we have

$$a(u, \phi_h^c) - a_h(v_h, \phi_h^c) = \sum_{K \in T_h} \int_K (a(u, \nabla u) - a(v_h, \nabla v_h)) \cdot \nabla \phi_h^c \, dx$$

$$+ \sum_{K \in T_h} \int_K (a_0(u, \nabla u) - a_0(v_h, \nabla v_h)) \phi_h^c \, dx.$$

Then, by (2.8), Cauchy-Schwarz inequality, (2.12) and Lemma 4.1

$$|a(u, \phi_h^c) - a_h(v_h, \phi_h^c)| \leq C|||u - v_h||| \cdot |||\phi_h^c|||_{1,h}$$

$$\leq C|||u - v_h||| \left( |||\phi_h^c|||_{1,h} + |||\phi_h^c|||_{1,h} \right)$$

$$\leq C|||u - v_h||| \cdot |||\phi_h|||$$

$$\leq C|||u - v_h||| \cdot |||u_h - v_h|||.$$

Obviously,

$$\frac{(f, \phi_h - \phi_h^c)}{|||\phi_h|||} - a_h(v_h, \varphi_h - \phi_h^c) \leq \sup_{\varphi \in V_h \setminus \{0\}} \frac{(f, \varphi - \varphi_h^c) - a_h(v_h, \varphi - \varphi_h^c)}{|||\varphi|||}$$
Then, (5.7) becomes

It follows from Green’s formula and (4.9) that

**Lemma 5.2.** There exists a positive constant $C$ independent of $h$ and $\gamma$ such that

$$
(5.6) \quad \frac{(f, \varphi_h - \varphi_h^c) - ah(v_h, \varphi_h - \varphi_h^c)}{|||\varphi_h|||} \leq C(|||u - v_h||| + \text{osc}_h(v_h)).
$$

**Proof.** Let $\psi_h = \varphi_h - \varphi_h^c$. From the definition of $a_h(\cdot, \cdot)$, we have

$$
(f, \psi_h) - a_h(v_h, \psi_h) = (f, \psi_h) - \sum_{K \in \mathcal{T}_h} \int_K \mathbf{a}(v_h, \nabla v_h) \cdot \nabla \psi_h \, dx
$$

$$
- \sum_{K \in \mathcal{T}_h} \int_K \mathbf{a}_0(v_h, \nabla v_h) \psi_h \, dx
$$

$$
+ \sum_{e \in \Gamma} \int_{e} \{\mathbf{a}(v_h, \nabla v_h)\} [\psi_h] \, ds - \sum_{e \in \Gamma} \frac{\gamma}{h_e} \int_{e} [v_h] [\psi_h] \, ds.
$$

(5.7)

It follows from Green’s formula and (4.9) that

$$
\sum_{K \in \mathcal{T}_h} \int_K \mathbf{a}(v_h, \nabla v_h) \cdot \nabla \psi_h \, dx
$$

$$
= - \sum_{K \in \mathcal{T}_h} \int_K \nabla \cdot \mathbf{a}(v_h, \nabla v_h) \psi_h \, dx + \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathbf{a}(v_h, \nabla v_h) \cdot \nu_K \psi_h \, ds
$$

$$
= - \sum_{K \in \mathcal{T}_h} \int_K \nabla \cdot \mathbf{a}(v_h, \nabla v_h) \psi_h \, dx + \sum_{e \in \Gamma} \int_{e} \{\mathbf{a}(v_h, \nabla v_h)\} [\psi_h] \, ds
$$

$$
+ \sum_{e \in \Gamma} \int_{e} [\mathbf{a}(v_h, \nabla v_h)] [\psi_h] \, ds.
$$

Then, (5.7) becomes

$$
(f, \psi_h) - a_h(v_h, \psi_h) = \sum_{K \in \mathcal{T}_h} \int_K (f + \nabla \cdot \mathbf{a}(v_h, \nabla v_h) - \mathbf{a}_0(v_h, \nabla v_h)) \psi_h \, dx
$$

$$
- \sum_{e \in \Gamma} \int_{e} [\mathbf{a}(v_h, \nabla v_h)] [\psi_h] \, ds - \sum_{e \in \Gamma} \frac{\gamma}{h_e} \int_{e} [v_h] [\psi_h] \, ds
$$

$$
= T_1 + T_2 + T_3.
$$

(5.8)

From the proof of Theorem 4.7, we know that

$$
(5.9) \quad \sum_{K \in \mathcal{T}_h} h_K^2 |||f + \nabla \cdot \mathbf{a}(v_h, \nabla v_h) - \mathbf{a}_0(v_h, \nabla v_h)|||_{0,K}^2 \leq C |||u - v_h|||_{1,h}^2 + \text{Cosc}_{\gamma}^2(v_h),
$$

and

$$
(5.10) \quad \sum_{e \in \Gamma} h_e |||\mathbf{a}(v_h, \nabla v_h)|||_{0,e}^2 \leq C |||u - v_h|||_{1,h}^2 + \text{Cosc}_{\gamma}^2(v_h).
$$

Applying Cauchy-Schwarz inequality, (5.9) and (4.4) gives
\[
|T_1| \leq \sum_{K \in T_h} \|f + \nabla \cdot a(v_h, \nabla v_h) - a_0(v_h, \nabla v_h)\|_{0,K} \|\psi_h\|_{0,K}
\]
(5.11) \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{K \in T_h} h_{K}^{-2} \|\varphi_h - \varphi_h^c\|_{0,K}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{K \in T_h} h_{K}^{-2} \|\psi_h\|_{0,K}^2 \right)^{1/2}
(5.12) \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{K \in T_h} h_{K}^{-2} \|\varphi_h - \varphi_h^c\|_{0,K}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \|\varphi_h\|_{\Gamma}
\]
The second term on the right-hand side of (5.8) can be estimated as $T_1$ by using (5.10), Lemma 2.2 and (4.4),
\[
|T_2| \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{\gamma \in \Gamma} h_{\gamma}^{-1} \|\psi_h\|_{0,\gamma}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{K \in T_h} h_{K}^{-2} \|\varphi_h - \varphi_h^c\|_{0,K}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{K \in T_h} h_{K}^{-2} \|\varphi_h\|_{\Gamma}^2 \right)^{1/2}
(5.13) \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{\gamma \in \Gamma} h_{\gamma}^{-1} \|\psi_h\|_{0,\gamma}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \|\varphi_h\|_{\Gamma}
\]
Since $u \in H^1_0(\Omega)$, we have $[u]_e = 0, \forall e \in \Gamma$. Then, the third term on the right-hand side of (5.8) becomes
\[
T_3 = -\sum_{\gamma \in \Gamma} h_{\gamma} \int_{\gamma} [v_h - u]\psi_h ds.
\]
Using Cauchy-Schwarz inequality, $[\varphi_h^c] = 0$ and Lemma 4.1, we get
\[
|T_3| \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{\gamma \in \Gamma} h_{\gamma}^{-1} \|\psi_h\|_{0,\gamma}^2 \right)^{1/2}
\leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{\gamma \in \Gamma} h_{\gamma}^{-1} \|\psi_h\|_{0,\gamma}^2 \right)^{1/2}
(5.14) \leq C(\|u - v_h\|_{1,h} + \text{osc}_h(v_h)) \left( \sum_{\gamma \in \Gamma} h_{\gamma}^{-1} \|\psi_h\|_{0,\gamma}^2 \right)^{1/2}
\]
Then, the desired result follows from (5.8), (5.11)-(5.13) and (2.12). □

From Lemmas 5.1 and 5.2, we have

**Theorem 5.3.** Assume that $u \in H^1(\Omega)$ and $u_h \in V_h$ are the solutions of (2.1) and (2.11), respectively. Then, there exists a positive constant $C$ independent of $h$ and $\gamma$ such that

\[
\|u - u_h\|_{1,h} \leq C \inf_{v_h \in V_h} (\|u - v_h\|_{1,h} + \text{osc}_h(v_h)).
\]

If $u \in H^{1+\epsilon}(\Omega), 0 < \epsilon \leq 1$, and using the discontinuous piecewise linear finite element space $V^1_h$ in (2.11), we have

**Theorem 5.4.** Assume that $u \in H^{1+\epsilon}(\Omega), 0 < \epsilon \leq 1$, and $u_h \in V^1_h$ are the
solutions of (2.1) and (2.11), respectively. Then, there exists a positive constant $C$ independent of $h$ such that
\begin{equation}
|||u - u_h||| \leq C h^2 |u|_{1+\varepsilon,2} + o(h).
\end{equation}

**Proof.** It follows from Theorem 5.3 that
\begin{equation}
|||u - u_h||| \leq C |||u - u_I||| + \text{osc}_h(u_I).
\end{equation}
Then, by the definition of the energy norm $||| \cdot |||$, Lemma 2.1 and the following interpolation estimates
\begin{align*}
|||u - u_I|||_{0,K} &\leq C h^{1+\varepsilon} |u|_{1+\varepsilon,2,K}, \quad |||u - u_I|||_{1,K} \leq C h^\varepsilon |u|_{1+\varepsilon,2,K},
\end{align*}
we get
\begin{align}
|||u - u_I|||^2 &= |||u - u_I|||_{0,h}^2 + \sum_{\gamma \in \Gamma} \frac{\gamma}{h_{\gamma}} \|u - u_I\|_{0,\partial \gamma}^2 \\
&\leq C h^{2\varepsilon} |u|_{1+\varepsilon,2}^2 + \sum_{K \in T_h} \frac{\gamma}{h_K} |||u - u_I|||_{0,\partial K}^2 \\
&\leq C h^{2\varepsilon} |u|_{1+\varepsilon,2}^2 + C \sum_{K \in T_h} (h_K^{-2} |||u - u_I|||_{0,K}^2 + ||u - u_I||_{1,K}^2) \\
&\leq C h^{2\varepsilon} |u|_{1+\varepsilon,2}^2.
\end{align}

From the definition of the total oscillation, we know that $\text{osc}_h(u_I)$ is a higher order term, which tends to zero faster that $O(h)$, i.e., $\text{osc}_h(u_I) = o(h)$, see [10, 54] for details. Then, the desired result follows from (5.16) and (5.17). \qed

**Acknowledgments**

**References**


Department of Mathematics, Yantai University, Shandong, 264005, People’s Republic of China. 
*E-mail*: bicj@ytu.edu.cn

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong. 
*E-mail*: yanping.lin@inet.polyu.edu.hk