A Posteriori Estimates of Taylor-Hood Element for Stokes Problem Using Auxiliary Subspace Techniques

Jiachuan Zhang1 · Ran Zhang2 · Xiaoshen Wang3

Received: 21 December 2021 / Revised: 10 May 2022 / Accepted: 29 June 2022 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Based on the auxiliary subspace techniques, a hierarchical basis a posteriori error estimator is proposed for the Stokes problem in two and three dimensions. For the error estimator, we need to solve only two global diagonal linear systems corresponding to the degree of freedom of velocity and pressure respectively, which reduces the computational cost sharply. The upper and lower bounds up to an oscillation term are shown without saturation assumption. Numerical simulations are performed to demonstrate the reliability of the a posteriori error estimator.

Keywords A posteriori error estimate · Taylor-Hood element · Auxiliary subspace techniques · Adaptive method · Stokes problem

Mathematics Subject Classification 65N15 · 65N30 · 65M12 · 76D07

1 Introduction

In this paper, we propose an a posteriori error estimator based on the auxiliary subspace techniques for the Taylor-Hood finite element method (FEM) [8, 9] to solve Stokes equations [18, 26] with Dirichlet boundary condition.

1 School of Physical and Mathematical Sciences, Nanjing Tech University, Nanjing 211816, People’s Republic of China
2 School of Mathematics, Jilin University, Changchun 130012, People’s Republic of China
3 Department of Mathematics and Statistics, University of Arkansas at Little Rock, Arkansas 72204, USA

Published online: 24 August 2022
where $\Omega \subset \mathbb{R}^d (d = 2, 3)$ is a bounded polygonal or polyhedral domain with the boundary $\Gamma$. The function $u$ is a vector velocity field and $p$ is the pressure. The functions $f$ and $g$ are given Lebesgue square-integrable functions on $\Omega$ and $\Gamma$, respectively. The problem (1.1)-(1.3) has a unique solution in the sense that $p$ is only determined up to an additive constant. In the later sections, we will analyze the case of $g = 0$, and the case $g \neq 0$ is similar.

A posteriori error estimators and adaptive FEM can be used to solve the problems with local singularities effectively. Hierarchical basis a posteriori estimator is a popular approach and has been proven to be robust and efficient, whose origins can be traced back to [27, 28]. In this approach, let $V_k$ and $W_{k+d}$ be the approximation space and auxiliary space, respectively, where $V_k \cap W_{k+d} = \{0\}$ (to be specified in Sect. 2). The solution of approximation problem (2.8) is denoted by $(\hat{u}, \hat{p}) \in V_k$. Then the approximation error $\| (u - \hat{u}, p - \hat{p}) \|_V$ can be estimated in auxiliary space $W_{k+d}$ with the help of the error problem (2.15). Traditionally, the upper and lower bounds of error estimations need to make use of a saturation assumption, i.e. the best approximation of $(u, p)$ in $V_k \cup W_{k+d}$ is strictly better than its best approximation in $V_k$. Although saturation assumption is widely accepted in a posteriori error analysis [1, 16] and satisfied in the case of small data oscillation [13], it is not difficult to construct counter-examples for particular problems on particular meshes [11]. To remove the saturation assumption, Araya et al. presented an adaptive stabilized FEM combined with a hierarchical basis a posteriori error estimator in a special auxiliary bubble function spaces for generalized Stokes problem and Navier-Stokes equations. The error analysis of upper and lower bounds avoids the use of saturation assumption. Although the construction of auxiliary space needs a transformation operator in the reference element, it provides a novel idea for removing saturation assumption in reliability analysis [2–5]. Hakula et al. constructed the auxiliary space directly on each element for the second order elliptic problem and elliptic eigenvalue problem and proved that the error is bounded by the error estimator up to oscillation terms without the saturation assumption [15, 17].

The contribution of this paper is twofold. Firstly, we extend the auxiliary subspace techniques in [17] to the Stokes problem in two and three dimensions. More specifically, we construct auxiliary spaces for velocity and pressure, respectively and prove that these auxiliary spaces satisfy the inf-sup condition shown in Lemma 2.2. The error $\| (u - \hat{u}, p - \hat{p}) \|_V$ can be bounded by the solution of the error problem (2.15) and the oscillation term $\text{osc}(f)$ (Theorem 3.1). We emphasize that the error analysis does not use the saturation assumption. The other contribution of the present work is the diagonalization of the error problem to reduce the computational cost. Considering that the Stokes problem is a saddle point problem, we replace part of the matrix, which is related to velocity only, with a diagonal matrix in (4.4) to construct the second error problem shown in (4.8). Then the solution of (4.8) combined with the oscillation term $\text{osc}(f)$ can be used to bound the error $\| (u - \hat{u}, p - \hat{p}) \|_V$ (Theorem 4.2). Here, obtaining the pressure and velocity requires solving a non-diagonal and diagonal linear system, respectively. To further reduce the computation, the diagonal matrix is obtained by multiplying the diagonal matrix of pressure correlation matrix by a constant $c_s$ related to the number of the bases of pressure in each element. Now, the linear systems of pressure and velocity are both diagonal, which is the third error problem shown in (4.25) whose solution combined with the oscillation term $\text{osc}(f)$ can be used to bound the error $\| (u - \hat{u}, p - \hat{p}) \|_V$ (Theorem 4.4).
The rest of the work is organized as follows. In Sect. 2, the FEM spaces, the approximation problem, and the first error problem are introduced. Section 3 presents a quasi-interpolant based on moment conditions and develops a posteriori error estimation related to the first error problem for the Stokes equation. In Sect. 4, to reduce the computational cost, the system diagonalization techniques are developed for velocity (the second error problem) and pressure (the third error problem), respectively. The a posteriori error estimates of the corresponding error problems are shown. In Sect. 5, we obtained the local and global a posteriori error estimators, and an adaptive FEM is proposed based on the solution of the three error problems. In Sect. 6, numerical experiment results are presented to verify the effectiveness of our a posteriori error estimators. The last section is devoted to some concluding remarks.

2 Approximation Problem and Error Problem

The following notations are used in this paper

\[ a(w, v) = \int_{\Omega} \nabla w : \nabla v, \]
\[ b(v, q) = \int_{\Omega} q \nabla \cdot v, \]
\[ a_1((w, r), (v, q)) = a(w, v) - b(v, r) + b(w, q), \]
\[ f(v) = \int_{\Omega} f \cdot v, \]

for all \((v, q), (w, r) \in V := [H^1_0(\Omega)]^d \times L^2_0(\Omega),\)

\[ H^1_0(\Omega) = \{ v \in H^1(\Omega) \mid v = 0 \text{ on } \Gamma \}, \]
\[ L^2_0(\Omega) = \{ q \in L^2(\Omega) \mid \int_{\Omega} q = 0 \}. \]

The variational formulation of (1.1)-(1.3) is: Find \((u, p) \in V\) such that

\[ a_1((u, p), (v, q)) = f(v), \quad (2.1) \]

for all \((v, q) \in V\).

We denote by \(\| \cdot \|_{m, \Omega}\) and \(| \cdot |_{m, \Omega}\) the standard norm and semi-norm of Sobolev space with \(m \geq 0\), respectively. For the sake of convenience, we will use \(\| \cdot \|\) and \(| \cdot |\) for \(\| \cdot \|_{0, \Omega}\) and \(| \cdot |_{0, \Omega}\), respectively. For the coupling space \(V\), we define

\[ \|(v, q)\|_V = \sqrt{\| \nabla v \|^2 + \| q \|^2}. \quad (2.2) \]

From Cauchy-Schwarz inequality, \(a_1(\cdot, \cdot)\) is continuous, i.e.

\[ |a_1((w, r), (v, q))| \leq C_1 \|(w, r)\|_V \|(v, q)\|_V. \quad (2.3) \]

From Proposition 4.69 in [26], \(a_1(\cdot, \cdot)\) satisfies the estimates

\[ \inf_{(v, q) \in V \setminus \{0\}} \sup_{(w, r) \in V \setminus \{0\}} \frac{a_1((v, q), (w, r))}{\|(v, q)\|_V \|(w, r)\|_V} \geq c_1. \quad (2.4) \]

We refer to \(C_1\) and \(c_1\) as the continuity and inf-sup constant, respectively.
2.1 Approximation Problem

Let $\mathcal{T}$ be a family of conforming, shape-regular simplicial partition of $\Omega$. Let $\mathcal{F}$ denote the set of $(d - 1)$-dimensional sub-simplices, the “faces” of $\mathcal{T}$, and further decompose it as $\mathcal{F} = \mathcal{F}_I \cup \mathcal{F}_D$, where $\mathcal{F}_I$ comprises those faces in the interior of $\Omega$, and $\mathcal{F}_D$ comprises those faces in $\Gamma$. To ensure that the Taylor-Hood element satisfies the stability condition (inf-sup condition), we make the following assumptions for $\mathcal{T}$:

Assumption 1. $\mathcal{T}$ contains at least three triangles in the case of $d = 2$.
Assumption 2. Every element $T \in \mathcal{T}$ has at least one vertex in the interior of $\Omega$ in the case of $d = 3$.

In our scheme, in order to have a conforming approximation we shall choose the finite-dimensional spaces $VV_{k+1}$ and $VP_k$ with $k \geq 1$ (called Hood-Taylor or Taylor-Hood element)

\[
VV_{k+1} = \{ \hat{v} \in [H^1_0(\Omega)]^d \mid \hat{v}|_T \in [P_{k+1}]^d, \forall T \in \mathcal{T} \} \subset [H^1_0(\Omega)]^d,
\]

\[
VP_k = \{ \hat{q} \in H^1(\Omega) \mid \hat{q}|_T \in P_k(K), \forall T \in \mathcal{T}, \int_{\Omega} \hat{q} = 0 \} \subset L^2_0(\Omega),
\]

\[
V_k = VV_{k+1} \times VP_k.
\]

A mixed finite element method to approximate (2.1) is called an approximation problem: Find $(\hat{u}, \hat{p}) \in V_k$ such that

\[
a_1((\hat{u}, \hat{p}), (\hat{v}, \hat{q})) = f(\hat{v}),
\]

for all $(\hat{v}, \hat{q}) \in V_k$.

Remark 2.1 The solvability of the approximation problem (2.8) can be found in [8–10].

2.2 Error Problem

Given a simplex $T \subset \mathbb{R}^d$ of diameter $h_T$, we define $S_j(T), 0 \leq j \leq d$ to be the set of sub-simplices of $T$ of dimension $j$. The cardinality is $|S_j(T)| = \binom{d+1}{j+1}$. We denote by $S_j$ the set of sub-simplices of the triangulation of dimension $j$, in particular, $S_{d-1} = \mathcal{F}_I \cup \mathcal{F}_D$ and $S_d = \mathcal{T}$. Recall that $P_m(S)$ is the set of polynomials of total degree $\leq m$ with domain $S$, and note that $\dim P_m(S) = \binom{m+j}{j}$ for $S \in S_j(T)$. Denoting the vertices of $T$ by $\{z_0, \ldots, z_d\}$, we let $\lambda_i \in P_1(T), 0 \leq i \leq d$, be the corresponding barycentric coordinates, uniquely defined by the relation $\lambda_i(z_j) = \delta_{ij}$. We denote by $F_j \in S_{d-1}(T)$ the sub-simplex not containing $z_j$.

The fundamental element and face bubbles for $T$ are given by ($j = 0, 1, \ldots, d$)

\[
b_T = \prod_{k=0}^{d} \lambda_k \in P_{d+1}(T), \quad b_{F_j} = \prod_{k=0}^{d} \lambda_k \in P_d(T).
\]

We also define general element and face bubbles of degree $m$,

\[
Q_m(T) = \{ \hat{v} = b_T \hat{w} \in P_m(T) \mid \hat{w} \in P_{m-d-1}(T) \},
\]

\[
Q_m(F_j) = \{ \hat{v} = b_{F_j} \hat{w} \in P_m(T) \mid \hat{w} \in P_{m-d}(T) \} \ominus Q_m(T).
\]

From now on, we use the shorthand $W_1 \ominus W_2 = \text{span}\{W_1 \setminus W_2\}$ for vector spaces $W_1$ and $W_2$. So $W_1 \ominus W_2$ is the largest subspace of $W_1$ such that $(W_1 \ominus W_2) \cap W_2 = \{0\}$. The functions
in \( Q_m(T) \) are precisely those in \( P_m(T) \) that vanish on \( \partial T \), and the functions in \( Q_m(F_j) \) are precisely those in \( P_m(T) \) that vanish on \( \partial T \setminus F_j \). It is clear that \( Q_m(T) \cap Q_m(F_j) = \{ 0 \} \) and \( Q_m(F_i) \cap Q_m(F_j) = \{ 0 \} \) for \( i \neq j \). The collection of face bubbles of degree \( m \) can be denoted by

\[
Q_m(\partial T) = \bigoplus_{j=0}^{d} Q_m(F_j).
\]

Then we define the local space

\[
R_m(T) = Q_m(T) \oplus Q_m(\partial T),
\]

which contains all element and face bubbles of degree \( m \) related to \( T \) defined in (2.9) and (2.10), and the corresponding global finite element spaces

\[
R_m = \{ \hat{v} \in H_0^1(\Omega) \mid \hat{v}\big|_T \in R_m(T) \text{ for each } T \in T \}.
\]

**Lemma 2.1** A function \( \hat{v} \in R_m(T) \) is uniquely determined by the moments

\[
\int_S \hat{v}_\kappa, \quad \forall \kappa \in P_{m-\ell-1}(S), \quad \forall S \in S_\ell(T), \quad d-1 \leq \ell \leq d.
\]

**Proof** As is shown in [6], a function \( v \in P_m(T) \) is uniquely determined by the moments

\[
\int_S \hat{v}_\kappa, \quad \forall \kappa \in P_{m-\ell-1}(S), \quad \forall S \in S_\ell(T), \quad 0 \leq \ell \leq d,
\]

where \( \int_S \hat{v}_\kappa \) with \( S \in S_0(T) \) is understood to be the evaluation of \( \hat{v} \) at the vertex \( S \). Since \( \hat{v} \in R_m(T) \) is uniquely determined by its moments on \( T \) and \( F_j, j = 0, \cdots, d \), the result is clear. \( \Box \)

Given \( k \in \mathbb{N} \), we define the local error space for velocity by element and face bubbles

\[
WV_{k+d+1}(T) = [R_{k+d+1}(T) \ominus R_{k+1}(T)]^d,
\]

and for pressure by element bubbles

\[
WP_{k+d}(T) = Q_{k+d}(T) \ominus Q_k(T).
\]

The velocity and pressure error spaces are constructed this way to satisfy the inf-sup condition shown in Lemma 2.3.

The corresponding global finite element spaces, defined by the degrees of freedom and local spaces, are given by

\[
WV_{k+d+1} = \{ \hat{\mathbf{w}} \in [H_0^1(\Omega)]^d \mid \hat{\mathbf{w}}\big|_T \in WV_{k+d+1}(T) \text{ for each } T \in T \},
\]

(2.12)

\[
WP_{k+d} = \{ \hat{p} \in L_0^2(\Omega) \cap H^1(\Omega) \mid \hat{p}\big|_T \in WP_{k+d}(T) \text{ for each } T \in T \},
\]

(2.13)

\[
W_{k+d} = WV_{k+d+1} \times WP_{k+d},
\]

(2.14)

where \( V_k \cap W_{k+d} = \{ 0 \} \). Then the **error problem** is: Find \((\hat{\mathbf{u}}, \hat{p}) \in W_{k+d}\) such that

\[
a_1((\hat{\mathbf{u}}, \hat{p}), (\hat{\mathbf{v}}, \hat{q})) = f(\hat{\mathbf{v}}) - a_1((\hat{\mathbf{u}}, \hat{p}), (\hat{\mathbf{v}}, \hat{q})),
\]

(2.15)

for any \((\hat{\mathbf{v}}, \hat{q}) \in W_{k+d}\).
2.3 Solvability of Error Problem

The error problem is stable (in the sense of inf-sup condition) in $W_{k+d}$ from the following Lemmas 2.2 and 2.3. Let $\overline{P}_k$ be the set of homogeneous polynomials of degree $k \geq 1$. Then we define

$$WV_{k+j+1} = W_{k+d+1} \cap [\overline{P}_{k+j+1}]^d, \quad WP_{k+j} = W_{k+d} \cap \overline{P}_{k+j},$$

where $1 \leq j \leq d$.

**Lemma 2.2** Under Assumptions 1 and 2, there exist positive constants $\mu_j$ ($1 \leq j \leq d$) independent of $h$ such that

$$\sup_{\hat{v} \in WV_{k+j+1}} \frac{b(\hat{v}, \hat{q})}{\| \hat{v} \|_{1,\Omega}} \geq \mu_j \| \hat{q} \|, \quad \forall \hat{q} \in \overline{P}_{k+j},$$

where $k \geq 1$.

**Proof** The proof can be found in Appendix A. \qed

**Lemma 2.3** There exists a positive constant $\mu$ independent of $h$ such that

$$\sup_{\hat{v} \in WV_{k+j+1}} \frac{b(\hat{v}, \hat{q})}{\| \hat{v} \|_{1,\Omega}} \geq \mu \| \hat{q} \|, \quad \forall \hat{q} \in WP_{k+d},$$

where $k \geq 1$.

**Proof** It follows from $WP_{k+j} \subset \overline{P}_{k+j}$ and Lemma 2.2 that

$$\sup_{\hat{v} \in WV_{k+j+1}} \frac{b(\hat{v}, \hat{q})}{\| \hat{v} \|_{1,\Omega}} \geq \mu_j \| \hat{q} \|, \quad \forall \hat{q} \in WP_{k+j},$$

for $k \geq 1$ and $1 \leq j \leq d$. Then set $\mu = \min_{1 \leq j \leq d} \mu_j$ and complete the proof from the facts

$$WV_{k+d+1} = \bigoplus_{j=1}^d WV_{k+j+1}, \quad WP_{k+d} = \bigoplus_{j=1}^d WP_{k+j}.$$

\qed

From Lemma 2.3, the proof of the following lemma is similar to that of Proposition 4.69 in [26]. For the sake of completeness, we give the proof here.

**Lemma 2.4** The bilinear form $a_1((\hat{v}, \hat{q}), (\hat{w}, \hat{r}))$ satisfies the estimate

$$\inf_{(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}} \sup_{(\hat{w}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{a_1((\hat{v}, \hat{q}), (\hat{w}, \hat{r}))}{\| (\hat{v}, \hat{q}) \|_V \| (\hat{w}, \hat{r}) \|_V} \geq \frac{\mu^2}{(1 + \mu)^2},$$

where $\mu$ is a constant defined in Lemma 2.3.

**Proof** Let $(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}$ be an arbitrary but fixed function. The definition of $a_1(\cdot, \cdot)$ immediately implies that

$$a_1((\hat{v}, \hat{q}), (\hat{v}, \hat{q})) = \| \nabla \hat{v} \|.$$
Due to Lemma 2.3, there is a velocity field \( \hat{\mathbf{w}}_\delta \in WV_{k+d+1} \) with \( \| \nabla \hat{\mathbf{w}}_\delta \| = 1 \) such that
\[
\int_{\Omega} \hat{\mathbf{q}} \nabla \cdot \hat{\mathbf{w}}_\delta \geq \mu \| \hat{\mathbf{q}} \|.
\]
We therefore obtain for every \( \delta > 0 \)
\[
a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{v}} - \delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, \hat{\mathbf{q}})) = a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{v}}, \hat{\mathbf{q}})) - \delta \| \hat{\mathbf{q}} \| a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{w}}_\delta, 0))
\]
\[
= \| \nabla \hat{\mathbf{v}} \|^2 - \delta \| \hat{\mathbf{q}} \| \int_{\Omega} \nabla \hat{\mathbf{v}} : \nabla \hat{\mathbf{w}}_\delta + \delta \| \hat{\mathbf{q}} \| \int_{\Omega} \hat{\mathbf{q}} \nabla \cdot \hat{\mathbf{w}}_\delta
\]
\[
\geq \| \nabla \hat{\mathbf{v}} \|^2 - \delta \| \nabla \hat{\mathbf{v}} \| \| \hat{\mathbf{q}} \| + \delta \mu \| \hat{\mathbf{q}} \|^2
\]
\[
\geq \left( 1 + \delta \right) \| \nabla \hat{\mathbf{v}} \|^2 + \frac{1}{2} \delta \mu \| \hat{\mathbf{q}} \|^2.
\]
The choice of \( \delta = \frac{2\mu}{1 + \mu^2} \) yields
\[
a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{v}} - \delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, \hat{\mathbf{q}})) \geq \frac{\mu^2}{1 + \mu^2} \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|^2.
\]
On the other hand, we have
\[
\| (\hat{\mathbf{v}} - \delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, \hat{\mathbf{q}}) \|_V \leq \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V + \| (\delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, 0) \|_V
\]
\[
= \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V + \| \delta \| \hat{\mathbf{q}} \| \| \nabla \hat{\mathbf{w}}_\delta \|
\]
\[
= \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V + \| \delta \| \hat{\mathbf{q}} \|
\]
\[
\leq (1 + \delta) \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V
\]
\[
= \frac{1 + \mu^2 + 2\mu}{1 + \mu^2} \| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V.
\]
Combining these estimates we arrive at
\[
\sup_{(\hat{\mathbf{w}}, \hat{\mathbf{r}}) \in W_{k+d} \setminus \{0\}} \frac{a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{w}}, \hat{\mathbf{r}}))}{\| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V \| (\hat{\mathbf{w}}, \hat{\mathbf{r}}) \|_V} \geq \frac{a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{v}} - \delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, \hat{\mathbf{q}}))}{\| (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \|_V \| (\hat{\mathbf{v}} - \delta \| \hat{\mathbf{q}} \| \hat{\mathbf{w}}_\delta, \hat{\mathbf{q}}) \|_V}
\]
\[
\geq \frac{\mu^2}{1 + \mu^2}.
\]
Since \( (\hat{\mathbf{v}}, \hat{\mathbf{q}}) \in W_{k+d} \setminus \{0\} \) was arbitrary, this completes the proof. \( \square \)

**Theorem 2.1** The error problem (2.15) has a unique solution.

**Proof** For the system (2.15), one can easily check that \( a_1(\cdot, \cdot) \) is a continuous bilinear form on \( W_{k+d} \times W_{k+d} \subset V \times V \) by (2.3) and satisfies the inf-sup condition by Lemma 2.4. In addition, \( f(\hat{\mathbf{v}}) - a_1((\hat{\mathbf{u}}, \hat{\mathbf{p}}), (\hat{\mathbf{v}}, \hat{\mathbf{q}})) \) is a continuous linear functional on \( W_{k+d} \) and the bilinear form \( a_1(\cdot, \cdot) \) satisfies
\[
a_1((\hat{\mathbf{v}}, \hat{\mathbf{q}}), (\hat{\mathbf{v}}, \hat{\mathbf{q}})) = \| \nabla \hat{\mathbf{v}} \|^2 \geq C \| \hat{\mathbf{v}} \|^2 > 0, \quad \hat{\mathbf{v}} \neq 0,
\]
by Poincare’s inequalities. So by Theorem 5.2.1 in [7], the scheme (2.15) has a unique solution. \( \square \)
3 A Posteriori Error Estimation

In this section, a quasi-interpolant based on moment conditions will be shown in Lemma 3.1, which is used to get the a posteriori error estimate shown in Theorem 3.1.

**Lemma 3.1** Given \( v \in [H^1(\Omega)]^d \), there exits a \( \hat{v} \in V V_{k+1} \) and \( \hat{w} \in W V_{k+d+1} \) such that

1. \( \int_T (v - \hat{v} - \hat{w}) \cdot \kappa = 0 \) for all \( \kappa \in [P_k(T)]^d \) and \( T \in \mathcal{T} \).
2. \( \int_F (v - \hat{v} - \hat{w}) \cdot \kappa = 0 \) for all \( \kappa \in [P_k(F)]^d \) and \( F \in \mathcal{F}_T \).
3. \( |v - \hat{v} - \hat{w}|_{m,T} \leq C_T h_T^{-m} |v|_{1,\Omega_T} \) for \( m = 0, 1 \), where \( \Omega_T \) is a local patch of elements containing \( T \).
4. \( |v - \hat{v} - \hat{w}|_{0,F} \leq C_T h_F^{1/2} |v|_{1,\Omega_F} \), where \( h_F \) is the diameter of \( F \in \mathcal{F} \), and \( \Omega_F = \Omega_T \) for some \( T \in \mathcal{T} \) with \( F \subset \partial T \).
5. \( |\hat{w}|_{1,T} \leq C_T |v|_{1,\Omega_T} \) for each \( T \in \mathcal{T} \).

where \( C_T \) depends only on the dimension \( d \), polynomial degree \( k \), and the shape-regularity of \( \mathcal{T} \).

**Proof** Since functions in \( R_{k+d+1}(T) \) are uniquely determined by the moments (2.11), for \( m = 0, 1 \) the function \( \langle \cdot \rangle_{m,T} : [R_{k+d+1}(T)]^d \to \mathbb{R}_+ \) defined by

\[
\langle \phi \rangle_{m,T} = \max_{S \in S(T)} \sup_{d-1 \leq \ell \leq d} \frac{h_T^{d/2-\ell/2-2m}}{\| \kappa \|_{0,S}} \int_S \phi \cdot \kappa
\]

is a norm on \([R_{k+d+1}(T)]^d\).

Let \( \tilde{T} = \{ y = h_T^{-1} x : x \in T \} \), and for each \( \psi : T \to \mathbb{R} \), define \( \tilde{\psi} : \tilde{T} \to \mathbb{R} \) by \( \tilde{\psi}(y) = \psi(h_T y) \). Analogous definitions are given for the sub-simplices of \( T \) and \( \tilde{T} \) and functions defined on them. It is clear that \( |\phi|_{m,T} = h_T^{d/2-2m} |\tilde{\phi}|_{m,\tilde{T}} \), where \( | \cdot |_{0,T} = \| \cdot \|_{0,T} \). We also have for any \( S \in S(T) \)

\[
\frac{h_T^{d/2-\ell/2-2m}}{\| \kappa \|_{0,S}} \int_S \phi \cdot \kappa = \frac{h_T^{d/2-\ell/2-2m}}{\| \kappa \|_{0,\tilde{S}}} \int_{\tilde{S}} \tilde{\phi} \cdot \tilde{\kappa}
\]

Since \( h_{\tilde{T}} = 1 \), we set that \( \langle \phi \rangle_{m,T} = h_T^{d/2-2m} |\tilde{\phi}|_{m,T} \). Therefore there exists a scale-invariant constant \( C_T > 0 \) that depends solely on \( k, d, \) and \( m \) such that

\[
|\phi|_{m,T} = h_T^{d/2-2m} |\tilde{\phi}|_{m,\tilde{T}} \leq C_T h_T^{d/2-2m} \langle \phi \rangle_{m,T} = C_T \langle \phi \rangle_{m,T}. \tag{3.1}
\]

At this stage, we see that the local constant \( C_T \) in (3.1) may depend on the shape of \( T \), but not its diameter. For the rest of the argument, we make a shape-regularity assumption on \( \mathcal{T} \).

Next, denote by \( \hat{v}_1 \in V V_{k+1} \) the Scott-Zhang interpolant of \( v \) satisfying [22]

\[
|v - \hat{v}_1|_{m,T} \leq C_T h_T^{1-m} |v|_{1,\Omega_T}, \quad m = 0, 1, \tag{3.2}
\]

\[
|v - \hat{v}_1|_{0,\partial T} \leq C_T h_T^{1/2} |v|_{1,\Omega_T}, \tag{3.3}
\]

on each \( T \in \mathcal{T} \). Set \( \hat{v}_2 \in [R_{k+d+1}]^d \) such that

\[
\int_{S} \hat{v}_2 \cdot \kappa = \int_{S} (v - \hat{v}_1) \cdot \kappa, \quad \forall \kappa \in [P_{k+d-\ell}(S)]^d, \quad \forall S \in S_\ell, \quad d - 1 \leq \ell \leq d.
\]
By (3.1)-(3.3) we get
\[
|\hat{v}|_{m,T} \leq C_T \max_{S \in S_\mathcal{T}(T)} \sup_{d-1 \leq \ell \leq d} h_T^{d/2-\ell/2-m} \|\kappa\|_{0,S} \int_S \hat{v} \cdot \kappa
\]
\[
= C_T \max_{S \in S_\mathcal{T}(T)} \sup_{d-1 \leq \ell \leq d} h_T^{d/2-\ell/2-m} \|\kappa\|_{0,S} \int_S (v - \hat{v}_1) \cdot \kappa
\]
\[
\leq C_T (h_T^{1/2-m} \|v - \hat{v}_1\|_{0,T} + h_T^{-m} \|v - \hat{v}_1\|_{0,T}) \leq C_T h_T^{1-m} |v|_{1,\Omega_T}.
\]

Uniquely decomposing \(\hat{v}_2\) as \(\hat{v}_2 = \hat{v}_3 + \hat{w}\) with \(\hat{v}_3 \in V_{k+1}\) and \(\hat{w} \in W_{k+d+1}\), and setting \(\hat{v} = \hat{v}_1 + \hat{v}_3\) so that \(\hat{v} + \hat{w} = \hat{v}_1 + \hat{v}_2\), we see that properties (1)-(2) clearly hold, and
\[
\|v - \hat{v} - \hat{w}\|_{m,T} \leq \|v - \hat{v}_1\|_{m,T} + \|\hat{v}_2\|_{m,T} \leq C_T h_T^{1-m} |v|_{1,\Omega_T}.
\]

Therefore by the standard trace inequalities and the shape regularity of the mesh, we also have on \(F \subset \partial T\)
\[
\|v - \hat{v} - \hat{w}\|_{0,F} \leq C_T (h_F^{-1/2} \|v - \hat{v} - \hat{w}\|_{0,T} + h_F^{1/2} \|v - \hat{v} - \hat{w}\|_{1,T}) \leq C_T h_F^{1/2} |v|_{1,F}.
\]

Hence, properties (3)-(4) are satisfied.

Finally, since \(V_{k+1}(T) \cap W_{k+d+1}(T) = \{0\}\), the strengthened Cauchy-Schwarz inequality [14] gives the existence of a constant \(\gamma \in (0, 1)\) such that
\[
\int_T \nabla \hat{w} \cdot \nabla \hat{v}_3 \leq \gamma |\hat{w}|_{1,T} |\hat{v}_3|_{1,T}.
\]

Consequently, we have
\[
|\hat{v}_2|_{1,T}^2 = |\hat{w}|_{1,T}^2 + |\hat{v}_3|_{1,T}^2 + 2 \int_T \nabla \hat{w} : \nabla \hat{v}_3
\]
\[
\geq |\hat{w}|_{1,T}^2 + |\hat{v}_3|_{1,T}^2 - 2 \gamma |\hat{w}|_{1,T} |\hat{v}_3|_{1,T} \geq (1 - \gamma^2) |\hat{w}|_{1,T}^2.
\]

Therefore we find \(|\hat{v}_2|_{1,T} \leq \sqrt{(1 - \gamma^2)^{-1}} |\hat{w}|_{1,T} \leq C_T |v|_{1,\Omega_T} \).

For \((v, q) \in V, \) we have
\[
a_1((u - \hat{u}, p - \hat{p}), (v, q)) = f(v) - a_1(\hat{u}, \hat{p}, (v, q)), \quad (3.4)
\]
where \((u, p)\) and \((\hat{u}, \hat{p})\) are the solutions of (2.1) and (2.8), respectively. So,
\[
a_1((u - \hat{u}, p - \hat{p}), (v, q)) = \sum_{T \in T} \int_T (f \cdot v - \nabla \hat{u} : \nabla v + \nabla \cdot \hat{p} - \nabla \cdot \hat{u} q)
\]
\[
= \sum_{T \in T} \int_T (f \cdot v - (-\Delta \hat{u} \cdot v + \nabla \hat{p} \cdot v - \nabla \cdot \hat{u} q))
\]+ \sum_{T \in T} \int_{\partial T} (-\nabla \hat{u} \cdot n_T \cdot v + \hat{p} v \cdot n_T)

Lemma 3.2 For any \((v, q) \in V, (\hat{w}, \hat{r}) \in W_{k+d}, and (\hat{v}, \hat{q}) \in V_k, it holds that
\[
a_1((u - \hat{u}, p - \hat{p}), (v, q)) = a_1((\hat{u}, \hat{p}), (\hat{v}, \hat{q})) + \mathcal{R}(v - \hat{v}, q - \hat{q}) \quad (3.5)
\]
where \((u, p)\) and \((\hat{u}, \hat{p})\) are the solutions of (2.1) and (2.8), respectively, and
\[
\mathcal{R}(w, r) = f(w) - a_1((\hat{u}, \hat{p}), (w, r))
\]
\[
\sum_{T \in \mathcal{T}} \int_{\Omega} \left( ( f - R_T ) \cdot w + \nabla \cdot \hat{u} r \right) + \sum_{F \in \mathcal{F}} \int_{F} r_F \cdot w,
\]

for any \((w, r) \in [H^1_0(\Omega)]^d \times L^2_0(\Omega)\) and

\[
R_T = (-\Delta \hat{u} + \nabla \hat{p})|_T, \\
r_F = (-\nabla \hat{u} \cdot n_T + \hat{p} n_T)|_T - (-\nabla \hat{u} \cdot n_{T'} + \hat{p} n_{T'})|_{T'}.
\]

Here, \(T\) and \(T'\) are the simplices sharing the face \(F\), and \(n_T\) and \(n_{T'}\) are their outward unit normals.

**Proof** From (2.8), (2.15), and (3.4), we obtain

\[
\mathcal{R}(v - \hat{w} - \hat{v}, q - \hat{r} - \hat{q}) = \mathcal{R}(v, q) - \mathcal{R}(\hat{w}, \hat{r}) - \mathcal{R}(\hat{v}, \hat{q}) = f(v) - a_1((\hat{u}, \hat{p}), (v, q)) - (f(\hat{w}) - a_1((\hat{u}, \hat{p}), (\hat{w}, \hat{r}))) - (f(\hat{v}) - a_1((\hat{u}, \hat{p}), (\hat{v}, \hat{q}))) = a_1((u - \hat{u}, p - \hat{p}), (v, q)) - a_1((\hat{e}_u, \hat{e}_p), (\hat{w}, \hat{r})),
\]

which completes the proof. \(\square\)

We define the local oscillation for each \(T \in \mathcal{T}\) by

\[
osc(f, T)^2 = h_T^2 \inf_{\kappa \in [P_k(T)]^d} \|f - \kappa\|^2_{0,T}.
\]

Then define

\[
osc(f)^2 = \sum_{T \in \mathcal{T}} osc(f, T)^2.\tag{3.6}
\]

**Theorem 3.1** Let \((u, p), (\hat{u}, \hat{p}),\) and \((\hat{e}_u, \hat{e}_p)\) be the solutions of (2.1), (2.8), and (2.15), respectively. There are constants \(\hat{C}_* = \frac{\mu^2}{2\hat{C}_1(1+\mu)}\) and \(\hat{C}_* = \frac{c_1}{\epsilon_1}\) such that

\[
\hat{C}_* \|(\hat{e}_u, \hat{e}_p)\|_V \leq \|(u - \hat{u}, p - \hat{p})\|_V \leq \hat{C}_* \|(\hat{e}_u, \hat{e}_p)\|_V + \frac{C_T}{\epsilon_1} osc(f),\tag{3.7}
\]

where constants \(\hat{C}_1, c_1, \mu,\) and \(C_T\) are defined in (2.3), (2.4), (2.16), and Lemma 3.1.

**Proof** Given \(q \in L^2_0(\Omega)\) and \(\hat{r} = 0\), let \(\hat{q}\) be the \(L^2\)-projection of \(q\) onto \(P_k(T)\). It is clear that \(\hat{q} \in V P_k \subset L^2_0(\Omega)\) since \(\int_{\Omega} \hat{q} = \int_{\Omega} q = 0\). Then combining Lemmas 3.1, 3.2, and noting \(R_T \in [P_k(T)]^2\), \(r_F \in [P_k(F)]^2\), we determine that

\[\square\]
\[
|a_1((\mathbf{u} - \hat{\mathbf{u}}, p - \hat{\rho}), (\mathbf{v}, q))| \\
\leq |a_1((\hat{\mathbf{e}}, \hat{e}), (\hat{\mathbf{w}}, \hat{r}))| + \sum_{T \in \mathcal{T}} \|\mathbf{v} - \hat{\mathbf{v}}\|_{0,T} \inf_{\kappa \in [P_k(T)]^d} \|\mathbf{f} - \kappa\|_{0,T} \\
+ \sum_{T \in \mathcal{T}} \|\mathbf{v} - \hat{\mathbf{v}} - \hat{\mathbf{w}}\|_{0,T} \inf_{\kappa \in [P_k(T)]^d} \|\mathbf{R}_T - \kappa\|_{0,T} \\
+ \left| \sum_{T \in \mathcal{T}} \int_T (q - \hat{q} - \hat{r}) \nabla \cdot \hat{\mathbf{u}} \right| + \sum_{F \in \mathcal{F}_T} \|\mathbf{v} - \hat{\mathbf{v}}\|_{0,F} \inf_{\kappa \in [P_{k+1}(F)]^d} \|\mathbf{r}_F - \kappa\|_{0,F} \\
\leq C_1 \|\hat{\mathbf{e}}, \hat{e}\|_V \|\hat{\mathbf{w}}, \hat{r}\|_V + C_T \sum_{T \in \mathcal{T}} h_T \|\mathbf{v}\|_{1,\Omega_T} \inf_{\kappa \in [P_k(T)]^d} \|\mathbf{f} - \kappa\|_{0,T} \\
\leq C_1 \|\hat{\mathbf{e}}, \hat{e}\|_V \|\mathbf{v}, q\|_V + C_T \text{osc}(f) \|\mathbf{v}, q\|_V,
\]
for any \(\hat{\mathbf{w}} \in W_{V_{k+d+1}}\) and \(\hat{\mathbf{v}} \in V_{V_{k+1}}\). Then the right inequality of (3.7) follows from the inf-sup condition (2.4) of continuous problem:

\[
c_1 \|\mathbf{u} - \hat{\mathbf{u}}, p - \hat{\rho}\|_V \leq \sup_{(\mathbf{w}, r) \in \mathcal{V} \setminus \{0\}} \frac{a_1((\mathbf{u} - \hat{\mathbf{u}}, p - \hat{\rho}), (\mathbf{w}, r))}{\|\mathbf{w}, r\|_V}.
\]

From (2.17), (2.15), and (2.1),

\[
\frac{\mu^2}{(1 + \mu)^2} \|\hat{\mathbf{e}}, \hat{e}\|_V \leq \sup_{(\hat{\mathbf{w}}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{a_1((\hat{\mathbf{e}}, \hat{e}), (\hat{\mathbf{w}}, \hat{r}))}{\|\hat{\mathbf{w}}, \hat{r}\|_V} \\
= \sup_{(\hat{\mathbf{w}}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{f(\hat{\mathbf{w}}) - a_1((\hat{\mathbf{u}}, \hat{\rho}), (\hat{\mathbf{w}}, \hat{r}))}{\|\hat{\mathbf{w}}, \hat{r}\|_V} \\
= \sup_{(\hat{\mathbf{w}}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{a_1((\mathbf{u} - \hat{\mathbf{u}}, p - \hat{\rho}), (\hat{\mathbf{w}}, \hat{r}))}{\|\hat{\mathbf{w}}, \hat{r}\|_V} \\
\leq C_1 \|\mathbf{u} - \hat{\mathbf{u}}, p - \hat{\rho}\|_V.
\]

Then we get the left inequality of (3.7). \(\square\)

## 4 System Diagonalization

As stated, the computation of \((\hat{\mathbf{e}}, \hat{e})\) requires the formation and solution of a global system, so one might naturally be concerned that this approach is too expensive for practical consideration. Generally speaking, the hierarchical basis for \(W_{k+d}\) is typically made up of highly oscillatory functions with compact support, therefore we may approximate the stiffness matrix by a diagonal matrix, which reduces the cost of computation.

### 4.1 Diagonalization with Respect to Velocity

Let \(\{\phi_j\}_{j=1}^N\) be the bases for \(W_{k+d}\), i.e.

\[
W_{k+d} = \text{span}\{\phi_j\}_{j=1}^N.
\]

Let \(\{\varphi_j\}_{j=1}^{N_v}\) and \(\{\psi_j\}_{j=1}^{N_p}\) be the bases in \(W_{k+d}\) for velocity and pressure, respectively. It is clear that \(N = N_v + N_p\) and \(\{\phi_j\}_{j=1}^N = \{\varphi_j\}_{j=1}^{N_v} \cup \{\psi_j\}_{j=1}^{N_p}\).
Define an \( N_v \times N_v \) matrix \( A \) by \( A_{\ell, j} = a(\varphi_j, \varphi_\ell) \) and an \( N_v \times N_p \) matrix \( B \) by \( B_{\ell, j} = -b(\psi_j, \varphi_\ell) \). Then we can rewrite (2.15) in a matrix form

\[
\begin{bmatrix}
A & B \\
-B^T & 0
\end{bmatrix}
\begin{bmatrix}
x_u \\
x_p
\end{bmatrix}
= \begin{bmatrix}
F_v \\
F_p
\end{bmatrix},
\]

(4.1)

where \( x_u \) and \( x_p \) are the coefficients of \( \hat{e}_u \) and \( \hat{e}_p \) with respect to the bases, respectively; \( F_v \) and \( F_p \) are the vectors formed by the right-hand function of (2.15) acting on the bases of velocity and pressure, respectively. For any \((\hat{v}, \hat{q}) = \sum_{j=1}^N x_j \phi_j, (\hat{w}, \hat{r}) = \sum_{j=1}^N y_j \phi_j \in W_{k+d}\), we have

\[
a_1((\hat{v}, \hat{q}), (\hat{w}, \hat{r})) = y^T M x,
\]

(4.2)

where

\[
x = (x_1, \ldots, x_N)^T, \quad y = (y_1, \ldots, y_N)^T, \quad \text{and} \quad M = \begin{bmatrix} A & B \\ -B^T & 0 \end{bmatrix}.
\]

(4.3)

Let \( x_v \) be a vector composed of elements related to velocity in \( x \), then it holds

\[
\| (\hat{v}, \hat{q}) \|_V^2 = |\hat{v}|^2_{1, \Omega} + \| \hat{q} \|^2 = x_v^T A x_v + \| \hat{q} \|^2.
\]

Let \( D_v \) be the diagonal matrix with the same diagonal as \( A \) and \( M_v \) be

\[
M_v = \begin{bmatrix} D_v & B \\ -B^T & 0 \end{bmatrix}.
\]

(4.4)

Define

\[
a_2((\hat{v}, \hat{q}), (\hat{w}, \hat{r})) = y^T M_v x
\]

(4.5)

and norms

\[
\| (\hat{v}, \hat{q}) \|_D^2 = x_v^T D_v x_v + \| \hat{q} \|^2,
\]

(4.6)

\[
|\hat{v}|_D^2 = x_v^T D_v x_v.
\]

(4.7)

Now, we are at the stage to present the second error problem: Find \((\hat{e}_u, \hat{e}_p) \in W_{k+d}\) such that

\[
a_2(\hat{e}_u, \hat{e}_p, (\hat{v}, \hat{q})) = f(\hat{v}) - a_1((\hat{u}, \hat{p}), (\hat{v}, \hat{q})), \quad \forall (\hat{v}, \hat{q}) \in W_{k+d},
\]

(4.8)

where \( a_2(\cdot, \cdot) \) is specified in (4.5).

For any \( T \in \mathcal{T} \) and \((\hat{v}, \hat{q}) \in W_{k+d}\), denote by \( \{\varphi_T, j\}_{j=1}^{N_{v,T}} \) the basis functions of velocity related to \( T \), then \( \hat{v}_T := \hat{v}|_T = \sum_{j=1}^{N_{v,T}} x_{T,j} \varphi_T, j \) with \( \{x_T, j\}_{j=1}^{N_{v,T}} \) being the coefficients. Let \( \hat{v}_{T, j} := x_{T,j} \varphi_T, j \), then \( \hat{v}_T = \sum_{j=1}^{N_{v,T}} \hat{v}_{T,j} \). We can rewrite \( |\hat{v}|_{1, \Omega} \) and \( |\hat{v}|_D \) as follows:

\[
|\hat{v}|_{1, \Omega}^2 = \sum_{T \in \mathcal{T}} |\hat{v}_T|_{1,T}^2,
\]

\[
|\hat{v}|_D^2 = \sum_{T \in \mathcal{T}} \sum_{j=1}^{N_{v,T}} |\hat{v}_{T,j}|_{1,T}^2.
\]
We define the local norm of $| \cdot |_D$ by

$$
| \hat{\mathbf{w}} |_{D,T} = \sqrt{\sum_{j=1}^{N_{v,T}} |\hat{\mathbf{w}}_{T,j}|_{1,T}^2},
$$

(4.9)

where $N_{v,T}$ is the number of basis functions of velocity in element $T$.

**Lemma 4.1** There exist two positive constants $\beta_1$ and $\beta_2$ independent of $h$ such that

$$
\beta_1 \leq \frac{|\hat{\mathbf{w}}|_{1,T}^2}{|\hat{\mathbf{w}}|_{D,T}^2} \leq \beta_2, \quad \beta_1 \leq \frac{|\hat{\mathbf{w}}|_{1,\Omega}^2}{|\hat{\mathbf{w}}|_{D}^2} \leq \beta_2,
$$

(4.10)

for all $T \in T$ and $\hat{\mathbf{w}} \in WV_{k+d+1}$.

**Proof** We claim that there exist two positive constants $\beta_{1T}$ and $\beta_{2T}$ independent of $h$ such that

$$
\beta_{1T} \sum_{j=1}^{N_{v,T}} |\hat{\mathbf{w}}_{T,j}|_{1,T}^2 \leq |\hat{\mathbf{w}}_{T}|_{1,T}^2 \leq \beta_{2T} \sum_{j=1}^{N_{v,T}} |\hat{\mathbf{w}}_{T,j}|_{1,T}^2, \quad T \in T_h.
$$

(4.11)

where $N_{v,T}$ is the number of basis functions of velocity in element $T$.

For the first inequality in (4.11), divide $\Lambda = \{j \in N^+ | 1 \leq j \leq N_{v,T}\}$ into two subsets $\Lambda_1 = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 \cap \Lambda_2 = \emptyset$. From Theorem 1 in [14], it gets that

$$
\left( \sum_{j_1 \in \Lambda_1} \nabla \hat{\mathbf{w}}_{T,j_1}, \sum_{j_2 \in \Lambda_2} \nabla \hat{\mathbf{w}}_{T,j_2} \right) \leq \gamma_{v,T} \sqrt{\sum_{j_1 \in \Lambda_1} |\hat{\mathbf{w}}_{T,j_1}|_{1,T}^2 \sum_{j_2 \in \Lambda_2} |\hat{\mathbf{w}}_{T,j_2}|_{1,T}^2},
$$

(4.12)

where $0 \leq \gamma_{v,T} < 1$ is independent of $h$. Using the strengthened Cauchy inequality (4.12) and Cauchy-Schwarz inequality, we deduce

$$
|\hat{\mathbf{w}}_{T}|_{1,T}^2 = \left| \sum_{j=1}^{N_{v,T}} \hat{\mathbf{w}}_{T,j} \right|_{1,T}^2 = \left( \sum_{j=1}^{N_{v,T}} \nabla \hat{\mathbf{w}}_{T,j}, \sum_{j=1}^{N_{v,T}} \nabla \hat{\mathbf{w}}_{T,j} \right)
$$

$$
\geq \left( \sum_{j=2}^{N_{v,T}} \hat{\mathbf{w}}_{T,j} \right)^2_{1,T} - 2 \gamma_{v,T} |\hat{\mathbf{w}}_{T,1}|_{1,T} \sum_{j=2}^{N_{v,T}} |\hat{\mathbf{w}}_{T,j}|_{1,T}^2.
$$

By a similar argument, we obtain

$$
|\hat{\mathbf{w}}_{T}|_{1,T}^2 = \left| \sum_{j=1}^{N_{v,T}} \hat{\mathbf{w}}_{T,j} \right|_{1,T}^2 \geq (1 - \gamma_{v,T})^j |\hat{\mathbf{w}}_{T,j}|_{1,T}^2 \geq (1 - \gamma_{v,T}) \sum_{j=1}^{N_{v,T}} |\hat{\mathbf{w}}_{T,j}|_{1,T}^2.
$$
which implies the first inequality in (4.11) with $\beta_{1T} = (1 - \gamma_{v,T})^{N_{v,T}}$.

The second inequality in (4.11) follows from the Cauchy-Schwarz inequality with $\beta_{2T} = N_{v,T}$. Therefore, the claim (4.11) holds. Summing up (4.11) overall $T \in T$ and noting

$$\frac{|\hat{\omega}|_W^2}{|\hat{\omega}|_D^2} = \frac{\sum_{T \in T} |\hat{\omega}|_{1,T}^2}{\sum_{T \in T} \sum_{j=1}^{N_{v,T}} |\hat{\omega}_{T,j}|^2},$$

we arrive at the conclusion (4.10) with $\beta_1 = \min_{T \in T} (1 - \gamma_{v,T})^{N_{v,T}}$ and $\beta_2 = \max_{T \in T} N_{v,T}$. \hfill \Box

**Lemma 4.2** For any $(\hat{\mathbf{v}}, \hat{q}), (\hat{\mathbf{w}}, \hat{r}) \in W_{k+d}$, we have

$$a_2((\hat{\mathbf{v}}, \hat{q}), (\hat{\mathbf{w}}, \hat{r})) \leq C_2 \| (\hat{\mathbf{w}}, \hat{r}) \|_V \| (\hat{\mathbf{v}}, \hat{q}) \|_V,$$

(4.13)

where $C_2$ is a positive constant.

**Proof** For any $(\hat{\mathbf{v}}, \hat{q}) = \sum_{j=1}^{N} x_j \phi_j$, $(\hat{\mathbf{w}}, \hat{r}) = \sum_{j=1}^{N} y_j \phi_j \in W_{k+d}$, define $\mathbf{x} = (x_1, \cdots, x_N)^T$ and $\mathbf{y} = (y_1, \cdots, y_N)^T$. Let $\mathbf{x}_v$ and $\mathbf{y}_v$ be vectors composed of elements related to velocity in $\mathbf{x}$ and $\mathbf{y}$, respectively. Similarly, Let $\mathbf{x}_p$ and $\mathbf{y}_p$ be vectors composed of elements related to pressure in $\mathbf{x}$ and $\mathbf{y}$, respectively. Then using (4.5)~(4.7), Cauchy-Schwarz inequality, and Lemma 4.1, we have

$$a_2((\hat{\mathbf{v}}, \hat{q}), (\hat{\mathbf{w}}, \hat{r})) = y^T M_v x = [y_v^T y_p^T] \begin{bmatrix} D_v & B^T \\ -B^T & 0 \end{bmatrix} \begin{bmatrix} x_v \\ x_p \end{bmatrix}$$

$$= y_v^T D_v x_v - y_p^T B^T x_v + y_v^T B x_p$$

$$\leq |\hat{\mathbf{v}}|_D |\hat{\mathbf{w}}|_D - (\nabla \cdot \hat{\mathbf{v}})^r + (\nabla \cdot \hat{\mathbf{w}})$$

$$\leq |\hat{\mathbf{v}}|_D |\hat{\mathbf{w}}|_D + \| \nabla \cdot \hat{\mathbf{v}} \| \| \hat{\mathbf{r}} \| + \| \nabla \cdot \hat{\mathbf{w}} \| \| \hat{\mathbf{q}} \|$$

$$\leq |\hat{\mathbf{v}}|_D |\hat{\mathbf{w}}|_D + \sqrt{d} |\hat{\mathbf{1}}_{\Omega}| \| \hat{\mathbf{r}} \| + \sqrt{d} |\hat{\mathbf{1}}_{\Omega}| \| \hat{\mathbf{q}} \|$$

$$\leq |\hat{\mathbf{v}}|_D |\hat{\mathbf{w}}|_D + \sqrt{d} \| \hat{\mathbf{1}}_{\Omega} \| \| \hat{\mathbf{r}} \| + \sqrt{d} \| \hat{\mathbf{1}}_{\Omega} \| \| \hat{\mathbf{q}} \|^2$$

$$\leq \sqrt{d} \| \hat{\mathbf{v}} \|^2 + \| \hat{\mathbf{w}} \|^2 + d^2 \beta_2 (\| \hat{\mathbf{r}} \|^2 + \| \hat{\mathbf{q}} \|^2)$$

$$\leq C_2 \| (\hat{\mathbf{v}}, \hat{q}) \|_D \| (\hat{\mathbf{w}}, \hat{r}) \|_D,$$

where $C_2 = \max (\sqrt{d}, d^{1/2})$. \hfill \Box

**Lemma 4.3** The bilinear form $a_2((\hat{\mathbf{v}}, \hat{q}), (\hat{\mathbf{w}}, \hat{r}))$ satisfies the estimate

$$\inf_{(\hat{\mathbf{v}}, \hat{q}) \in W_{k+d} \setminus \{0\}} \sup_{(\hat{\mathbf{w}}, \hat{r}) \in W_{k+d} \setminus \{0\}} a_2((\hat{\mathbf{v}}, \hat{q}), (\hat{\mathbf{w}}, \hat{r})) \| (\hat{\mathbf{v}}, \hat{q}) \|_D \| (\hat{\mathbf{w}}, \hat{r}) \|_D \geq \frac{(\mu \beta_1)^2}{(1 + \mu \beta_1)^2},$$

where $\mu$ and $\beta_1$ are the constants in Lemmas 2.3 and 4.1, respectively.

**Proof** To prove the inequality, we choose an arbitrary but fixed element $(\hat{\mathbf{v}}, \hat{q}) \in W_{k+d} \setminus \{0\}$. Due to Lemma 2.3, there is a velocity field $\hat{\mathbf{w}}_\hat{q} \in \mathcal{W} V_{k+d+1}$ with $|\hat{\mathbf{w}}_\hat{q}|_D = 1$ such that

$$\sum_{T \in T} \int_T \hat{q} \nabla \cdot \hat{\mathbf{w}}_\hat{q} dx = \int_\Omega \hat{q} \nabla \cdot \hat{\mathbf{w}}_\hat{q} dx \geq \mu \| \hat{q} \|.$$
By using Cauchy-Schwartz inequality, Lemmas 4.1, 2.3, and noting $|\hat{w}_q|_D = 1$, we therefore obtain for every $\delta > 0$,

$$a_2((\hat{v}, \hat{q}), (\hat{v} - \delta \|\hat{q}\| \hat{w}_q, \hat{q})) = a_2((\hat{v}, \hat{q}), (\hat{v}, \hat{q})) - \delta \|\hat{q}\|a_2((\hat{v}, \hat{q}), (\hat{w}_q, 0))$$

$$= |\hat{v}|_D^2 - \delta \|\hat{q}\| \|\hat{v}\| \sum_{T \in T} \int_T \hat{q} \nabla \cdot \hat{w}_q$$

$$\geq |\hat{v}|_D^2 - \delta |\hat{v}|_D \|\hat{q}\| + \delta \mu \|\hat{q}\|^2 |\hat{w}_q|_{1, \Omega}$$

$$\geq |\hat{v}|_D^2 - \delta |\hat{v}|_D \|\hat{q}\| + \delta \mu \beta_1 \|\hat{q}\|^2$$

$$\geq (1 - \frac{\delta}{2\mu \beta_1}) |\hat{v}|_D^2 + \frac{1}{2} \frac{\delta}{\mu \beta_1} |\hat{q}|^2,$$

where $\hat{x}_v = (x_1, x_2, \ldots, x_n)^T$ and $\hat{y}_v = (y_1, y_2, \ldots, y_n)^T$ are such that $\hat{v} = \sum_{j=1}^{N_x} x_j \varphi_j, \hat{w}_q = \sum_{j=1}^{N_q} y_j \varphi_j \in W_{k+d+1}$.

Similar to the proof in Lemma 2.4, the choice of $\delta = \frac{2\mu \beta_1}{1 + (\mu \beta_1)^2}$ yields

$$a_2((\hat{v}, \hat{q}), (\hat{v} - \delta \|\hat{q}\| \hat{w}_q, \hat{q})) \geq \frac{(\mu \beta_1)^2}{1 + (\mu \beta_1)^2} \|\hat{v}\|_D^2,$$

and

$$\|\hat{v} - \delta \|\hat{q}\| \hat{w}_q, \hat{q})\|_D \leq \frac{(1 + \mu \beta_1)^2}{1 + (\mu \beta_1)^2} \|\hat{v}\|_D.$$

Then we arrive at

$$\sup_{(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}} \frac{a_2((\hat{v}, \hat{q}), (\hat{v}, \hat{r})) \|\hat{v}\|_D \|\hat{w}_q, \hat{q})\|_D \geq \frac{a_2((\hat{v}, \hat{q}), (\hat{v} - \delta \|\hat{q}\| \hat{w}_q, \hat{q})) \|\hat{v}\|_D \|\hat{w}_q, \hat{q})\|_D \geq \frac{(\mu \beta_1)^2}{1 + (\mu \beta_1)^2}.$$

Since $(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}$ is arbitrary, this completes the proof.

Using Lemmas 4.2, 4.3, and a proof similar to that of Theorem 2.1, we have the following conclusion.

**Theorem 4.1** The finite element scheme (4.8) has a unique solution.
which implies the first inequality in (4.14).

Similarly, using (4.15) and Lemma 2.3, we have

\[
\frac{\mu^2}{(1 + \mu)^2} \| (\hat{e}_u, \hat{e}_p) \|_V \leq \sup_{(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}} \frac{a_1((\hat{e}_u, \hat{e}_p), (\hat{v}, \hat{q}))}{\| (\hat{v}, \hat{q}) \|_D} \left( \frac{a_2((\hat{e}_u, \hat{e}_p), (\hat{v}, \hat{q}))}{\| (\hat{v}, \hat{q}) \|_V} \right)
\]

which implies the second inequality in (4.14).

Combining Theorem 3.1 and Lemma 4.4, we obtain the following lower and upper bounds related to \( \| (\hat{e}_u, \hat{e}_p) \|_D \).

**Theorem 4.2** Let \((u, p), (\hat{u}, \hat{p})\), and \((\hat{e}_u, \hat{e}_p)\) be the solutions of (2.1), (2.15), and (4.8), respectively. There are constants \(\bar{C}_*\) and \(\bar{C}^*\) such that

\[
\bar{C}_* \| (\hat{e}_u, \hat{e}_p) \|_D \leq \| (u - \hat{u}, p - \hat{p}) \|_V \leq \bar{C}^* \| (\hat{e}_u, \hat{e}_p) \|_D + \frac{C_T}{c_1} \text{osc}(f),
\]

where \(\| \cdot \|_V\) and \(\| \cdot \|_D\) are defined in (2.2) and (4.6), respectively. The constants \(C_1, C_2, \mu, \beta_1, \beta_2\), and \(C_T\) are defined in (2.3), (4.13), (2.4), (2.16), (4.11), and Lemma 3.1, respectively.
4.2 Diagonalization with Respect to Pressure

Recall that \( \{\varphi_j\}_{j=1}^{N_v} \) and \( \{\psi_j\}_{j=1}^{N_p} \) are the bases in space \( W_{k+d} \) for velocity and pressure, respectively. For \( \hat{e}_u = \sum_{j=1}^{N_v} \tilde{x}_{u,j} \varphi_j \) and \( \hat{e}_p = \sum_{j=1}^{N_p} \tilde{x}_{p,j} \psi_j \), rewrite (4.8) in a matrix form

\[
\begin{bmatrix}
D_v \\ -B^T
\end{bmatrix}
\begin{bmatrix}
\tilde{x}_u \\ \tilde{x}_p
\end{bmatrix}
= 
\begin{bmatrix}
F_v \\ F_p
\end{bmatrix},
\]  
(4.17)

where

\[
\tilde{x}_u = (\tilde{x}_{u,1}, \tilde{x}_{u,2}, \cdots, \tilde{x}_{u,N_v})^T, \quad F_v = (F_{v,1}, F_{v,2}, \cdots, F_{v,N_v})^T,
\]
\[
\tilde{x}_p = (\tilde{x}_{p,1}, \tilde{x}_{p,2}, \cdots, \tilde{x}_{p,N_p})^T, \quad F_p = (F_{p,1}, F_{p,2}, \cdots, F_{p,N_p})^T.
\]

Here, \( F_{v,j} \) and \( F_{p,j} \) are defined by

\[
F_{v,j} = f(\varphi_j) - a_1((\hat{u}, \hat{p}), (\varphi_j, 0)), \quad j = 1, 2, \cdots, N_v,
\]
\[
F_{p,j} = -a_1((\hat{u}, \hat{p}), (0, \psi_j)), \quad j = 1, 2, \cdots, N_p.
\]

After a simple calculation, we have

\[
D_v \tilde{x}_u + B \tilde{x}_p = F_v,
\]
(4.20)
\[
B^T D_v^{-1} B \tilde{x}_p = F_p + B^T D_v^{-1} F_v.
\]
(4.21)

The inverse of the matrix \( D_v \) is easy to calculate because it is a diagonal matrix. If we get \( \tilde{x}_p \) by solving (4.21), \( \tilde{x}_u \) is easy to get by (4.20). Let \( D_p = \text{diag}(B^T D_v^{-1} B) \), which is the diagonal matrix with the same diagonal as \( B^T D_v^{-1} B \). Let \( c_s = \max_{T \in T} N_{p,T} \), which is the maximum number of basis functions of pressure for each element. Then replacing \( B^T D_v^{-1} B \) with \( c_s D_p \) in (4.21), we get

\[
D_v \tilde{x}_u + B \tilde{x}_p = F_v,
\]
(4.22)
\[
c_s D_p \tilde{x}_p = F_p + B^T D_v^{-1} F_v.
\]
(4.23)

where \( \tilde{x}_u = (\tilde{x}_{u,1}, \tilde{x}_{u,2}, \cdots, \tilde{x}_{u,N_v}) \) and \( \tilde{x}_p = (\tilde{x}_{p,1}, \tilde{x}_{p,2}, \cdots, \tilde{x}_{p,N_p}) \).

Equations (4.22) and (4.23) are equivalent to

\[
D_v \tilde{x}_u + B \tilde{x}_p = F_v,
\]
\[
- B^T \tilde{x}_u + (c_s D_p - B^T D_v^{-1} B) \tilde{x}_p = F_p,
\]

whose matrix form is

\[
\begin{bmatrix}
D_v \\ -B^T
\end{bmatrix}
\begin{bmatrix}
\tilde{x}_u \\ \tilde{x}_p
\end{bmatrix}
= 
\begin{bmatrix}
F_v \\ F_p
\end{bmatrix}.
\]

For any \( (\hat{v}, \hat{q}) = \sum_{j=1}^{N} x_j \varphi_j \) and \( (\hat{w}, \hat{r}) = \sum_{j=1}^{N} y_j \psi_j \) in \( W_{k+d} \), we define

\[
a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r})) = y^T M_{vp} x,
\]
(4.24)

where

\[
y = (y_1, \cdots, y_N)^T, \quad M_{vp} = \begin{bmatrix}
D_v \\ -B^T c_s D_p - B^T D_v^{-1} B
\end{bmatrix}, \quad \text{and} \quad x = (x_1, \cdots, x_N)^T.
\]
It is time to present the third error problem: Find \( \{ \tilde{e}_u, \tilde{e}_p \} \in W_{k+d} \) with \( \tilde{e}_u = \sum_{j=1}^{N_p} \tilde{x}_{u,j} \phi_j \) and \( \tilde{e}_p = \sum_{j=1}^{N_p} \tilde{x}_{p,j} \psi_j \) such that
\[
a_3((\tilde{e}_u, \tilde{e}_p), (\hat{u}, \hat{q})) = f(\hat{u}) - a_1((\hat{u}, \hat{p}), (\hat{u}, \hat{q})), \quad \forall (\hat{u}, \hat{q}) \in W_{k+d}.
\] (4.25)

**Remark 4.1** Equations (4.22) and (4.23) are equivalent to (4.25), but they are used in different ways. Obviously, (4.22) and (4.23) are easier to calculate. In Sect. 5, the global and local estimators will be generated from (4.22) and (4.23). However, (4.25) is essential in the proof of equivalence. Therefore, we use (4.22) and (4.23) for the numerical computation and (4.25) for the theoretical analysis.

Because matrix \( D_u \) and \( D_p \) are diagonal matrices in (4.22) and (4.23), the existence and uniqueness of finite element scheme (4.25) are obvious.

**Theorem 4.3** The finite element scheme (4.25) has a unique solution.

Next, we turn our attention to the discrete pressure space. Two new norms will be defined. We still use \( \{ \psi_j \}_{j=1}^{N_p} \) to denote the basis functions of pressure in \( W_{k+d} \). For \( \hat{q} = \sum_{j=1}^{N_p} x_p^j \psi_j \) and \( \hat{r} = \sum_{j=1}^{N_p} y_p^j \psi_j \), define two bilinear forms
\[
E_{31}(\hat{q}, \hat{r}) = y_p^T B^T D_{v}^{-1} B x_p, \quad E_{32}(\hat{q}, \hat{r}) = y_p^T D_p x_p,
\]
and norms
\[
\|\hat{q}\|_B = \sqrt{E_{31}(\hat{q}, \hat{q})}, \quad \|\hat{q}\|_p = \sqrt{E_{32}(\hat{q}, \hat{q})},
\]
where \( x_p = (x_{p,1}, \ldots, x_{p,N_p})^T \) and \( y_p = (y_{p,1}, \ldots, y_{p,N_p})^T \). The next two lemmas will establish some inequalities related to the three pressure norms \( \| \cdot \|_B, \| \cdot \|_p \), and \( \| \cdot \| \).

**Lemma 4.5** There exist two positive constants \( c_i \) and \( c_s \) independent of \( h \) such that
\[
c_i \leq \frac{\|\hat{q}\|_B^2}{\|\hat{q}\|_p^2} \leq c_s, \quad \forall \hat{q} \in W_{P_{k+d}},
\] (4.26)
where \( c_s \) is the same as in (4.23).

**Proof** For any \( T \in T \), denote by \( \{ \psi_{T,j} \}_{j=1}^{N_{P,T}} \) the basis functions of pressure related to \( T \), then
\[
\hat{q}_T := \hat{q}|_T = \sum_{j=1}^{N_{P,T}} x_{T,j} \psi_{T,j} \quad \text{with } \{ x_{T,j} \}_{j=1}^{N_{P,T}} \text{ being the coefficients.}
\]
Let \( \hat{q}_{T,j} := x_{T,j} \psi_{T,j} \), then
\[
\hat{q}_T = \sum_{j=1}^{N_{P,T}} \hat{q}_{T,j}.
\]
We claim that there exist two positive constants \( c_iT \) and \( c_sT \), independent of \( h \), such that
\[
c_iT \sum_{j=1}^{N_{P,T}} \|\hat{q}_{T,j}\|_B^2 \leq \|\hat{q}_T\|_B^2 \leq c_sT \sum_{j=1}^{N_{P,T}} \|\hat{q}_{T,j}\|_B^2, \quad T \in T.
\] (4.27)

For the first inequality in (4.27), devide \( \Lambda = \{ j \in N^+ \mid 1 \leq j \leq N_{P,T} \} \) into two subsets \( \Lambda = \Lambda_1 \cup \Lambda_2 \) with \( \Lambda_1 \cap \Lambda_2 = \emptyset \). From Theorem 1 in [14], it gets that
\[
E_{31} \left( \sum_{j \in \Lambda_1} \hat{q}_{T,j}, \sum_{j \in \Lambda_2} \hat{q}_{T,j} \right) \leq y_{p,T} \left\| \sum_{j \in \Lambda_1} \hat{q}_{T,j} \right\|_B \left\| \sum_{j \in \Lambda_2} \hat{q}_{T,j} \right\|_B,
\] (4.28)
where \( 0 \leq \gamma_{p,T} < 1 \) is independent of \( h \). Using the strengthened Cauchy inequality (4.28), we deduce
\[
\left\| \sum_{j=1}^{N_{p,T}} \hat{q}_{T,j} \right\|_B^2 = \left\| \hat{q}_T \right\|_B^2 = E_{31} \left( \sum_{j=1}^{N_{p,T}} \hat{q}_{T,j}, \sum_{j=1}^{N_{p,T}} \hat{q}_{T,j} \right)
\]
\[
= \left\| \hat{q}_{T,1} \right\|_B^2 + \sum_{j=2}^{N_{p,T}} \left\| \hat{q}_{T,j} \right\|_B^2 + 2E_{31} \left( \hat{q}_{T,1}, \sum_{j=2}^{N_{p,T}} \hat{q}_{T,j} \right)
\]
\[
\geq \left\| \hat{q}_{T,1} \right\|_B^2 + \sum_{j=2}^{N_{p,T}} \left\| \hat{q}_{T,j} \right\|_B^2 - 2\gamma_{p,T} \left\| \hat{q}_{T,1} \right\|_B \left\| \sum_{j=2}^{N_{p,T}} \hat{q}_{T,j} \right\|_B
\]
\[
\geq (1 - \gamma_{p,T}) \left\| \hat{q}_{T,1} \right\|_B^2 + (1 - \gamma_{p,T}) \left\| \sum_{j=2}^{N_{p,T}} \hat{q}_{T,j} \right\|_B^2.
\]
By a similar argument, we obtain
\[
\left\| \hat{q}_T \right\|_B^2 = \left\| \sum_{j=1}^{N_{p,T}} \hat{q}_{T,j} \right\|_B^2 \geq \sum_{j=1}^{N_{p,T}} (1 - \gamma_{p,T})^j \left\| \hat{q}_{T,j} \right\|_B^2 \geq (1 - \gamma_{p,T})^{N_{p,T}} \sum_{j=1}^{N_{p,T}} \left\| \hat{q}_{T,j} \right\|_B^2,
\]
which implies the first inequality in (4.27) with \( c_{i,T} = (1 - \gamma_{p,T})^{N_{p,T}} \).

The second inequality in (4.11) follows from the Cauchy-Schwarz inequality with \( c_{i,T} = N_{p,T} \). Therefore, the claim (4.11) holds. Summing up (4.11) over all \( T \in T \) and noting
\[
\left\| \hat{q} \right\|_B^2 = \sum_{T \in T} \left\| \hat{q}_T \right\|_B^2 = \sum_{T \in T} \sum_{j=1}^{N_{p,T}} \left\| \hat{q}_{T,j} \right\|_B^2,
\]
we arrive at the conclusion (4.10) with \( c_i = \min_{T \in T} (1 - \gamma_{p,T})^{N_{p,T}} \) and \( c_s = \max_{T \in T} N_{p,T} \).  \( \Box \)

**Lemma 4.6** For any \( (\tilde{\psi}, \tilde{q}) \in W_{k+d} \), we have
\[
\left\| \tilde{q} \right\|_B \leq d(d + 1)\beta_2 \left\| \hat{q} \right\|
\]
where \( d \) is the dimension and \( \beta_2 \) is defined in (4.10).

**Proof** We continue to use \( \{\varphi_j\}_{j=1}^{N_v} \) and \( \{\psi_j\}_{j=1}^{N_p} \) as the bases in \( W_{k+d} \) for velocity and pressure, respectively. Define a diagonal matrix \( \tilde{D} \) whose elements are the square roots of the corresponding elements of \( D_v^{-1} \), and it is clear that \( D_v^{-1} = \tilde{D} \tilde{D} \). Let \( q = \sum_{j=1}^{N_p} x_j \psi_j \) and \( x = (x_1, x_2, \cdots, x_{N_p})^T \), then
\[
\left\| \hat{q} \right\|_B^2 = x^T B^T D_v^{-1} B x = x^T B^T \tilde{D} \tilde{D} B x.
\]
(4.29)

Let \( \{d_j\}_{j=1}^{N_v} \) denote the diagonal elements of the matrix \( \tilde{D} \) whose dimension is \( N_v \). Let \( d_j \) denote the \( j \)-th column of matrix \( \tilde{D} \) and \( \tilde{v}_j = d_j \varphi_j \) and denoted by \( T_1, \cdots, T_{j_T} \) the elements related to \( \varphi_j \) respectively. Then
\[
\sum_{i=1}^{j_T} \left\| \tilde{v}_j \right\|_{D,T_i}^2 = \left\| \tilde{v}_j \right\|_{D}^2 = d_j^T D_v d_j = 1.
\]
(4.30)
Lemma 4.7  The bi-linear form $a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r}))$ satisfies the estimates

$\inf_{(\hat{v}, \hat{q}) \in W_{k+d}^\perp \setminus \{0\}} \sup_{(\hat{w}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r}))}{\|\hat{v}\|_D \|\hat{w}\|_D} \geq \frac{(\mu \beta_1)^2}{(1 + \mu \beta_1)^2},$

where $\mu$ and $\beta_1$ are the constants in Lemmas 2.3 and 4.1, respectively.

Proof  To prove the inequality, we choose an arbitrary but fixed element $(\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}$. Due to Lemma 2.3, there is a velocity field $\hat{w}_\delta \in W_{k+d}$ with $|\hat{w}_\delta|_D = 1$ such that

$$\sum_{T \in \mathcal{T}} \int_T \hat{q} \nabla \cdot \hat{w}_\delta = \int_{\Omega} \hat{q} \nabla \cdot \hat{w}_\delta \geq \mu \|\hat{q}\|.$$

By using Cauchy-Schwartz inequality, Lemmas 4.1, 2.3, and 4.5, we therefore obtain for every $\delta > 0$

$$a_3((\hat{v}, \hat{q}), (\hat{w} - \delta \|\hat{q}\| \hat{w}_\delta, \hat{q}))$$

$$= a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{q})) - \delta \|\hat{q}\| a_3((\hat{v}, \hat{q}), (\hat{w}_\delta, 0))$$

$$= |\hat{v}|_D^2 + c_s \|\hat{q}\|_P^2 - |\hat{q}|_D^2 - \delta \|\hat{q}\| y_v D_v x_v + \delta \|\hat{q}\| \sum_{T \in \mathcal{T}_h} \int_T \hat{q} \nabla \cdot \hat{w}_\delta$$

$$\geq |\hat{v}|_D^2 - \delta |\hat{v}|_D \|\hat{q}\| + \delta \mu \|\hat{q}\|_D \|\hat{w}_\delta\|_{1, \Omega}$$

$$\geq |\hat{v}|_D^2 - \delta |\hat{v}|_D \|\hat{q}\| + \delta \mu \beta_1 \|\hat{q}\|_D$$

$$\geq \left(1 - \frac{\delta}{2 \mu \beta_1}\right) |\hat{v}|_D^2 + \frac{1}{2} \delta \mu \beta_1 \|\hat{q}\|_D^2,$$

where $x_v = (x_{v,1}, \ldots, x_{v,N_v})^T$, $x_q = (x_{q,1}, \ldots, x_{q,N_q})^T$, and $y_v = (y_{v,1}, \ldots, y_{v,N_v})^T$. Let $\hat{v} = \sum_{j=1}^{N_v} x_{v,j} \psi_j$, $\hat{q} = \sum_{j=1}^{N_q} x_{q,j} \psi_j$, and $\hat{w}_\delta = \sum_{j=1}^{N_v} y_{v,j} \psi_j$.

Similar to the proof of Lemma 2.4, the choice of $\delta = \frac{2 \mu \beta_1}{1 + (\mu \beta_1)^2}$ yields

$$a_3((\hat{v}, \hat{q}), (\hat{w} - \delta \|\hat{q}\| \hat{w}_\delta, \hat{q})) \geq \frac{(\mu \beta_1)^2}{1 + (\mu \beta_1)^2} \|\hat{v}\|_D^2.$$
and
\[ \| (\hat{v} - \delta \| \hat{q} \| \hat{w}_q, \hat{q}) \|_D \leq \frac{(1 + \mu \beta_1)^2}{1 + (\mu \beta_1)^2} \| (\hat{v}, \hat{q}) \|_D. \]

Then we arrive at
\[ \sup_{(\hat{w}, \hat{r}) \in W_{k+d} \setminus \{0\}} \frac{a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r}))}{\| (\hat{v}, \hat{q}) \|_D \| (\hat{w}, \hat{r}) \|_D} \geq \frac{a_3((\hat{v}, \hat{q}), (\hat{v} - \delta \| \hat{q} \| \hat{w}_q, \hat{q}))}{\| (\hat{v}, \hat{q}) \|_D \| (\hat{v} - \delta \| \hat{q} \| \hat{w}_q, \hat{q}) \|_D} \geq \frac{(\mu \beta_1)^2}{(1 + \mu \beta_1)^2}. \]

Since \((\hat{v}, \hat{q}) \in W_{k+d} \setminus \{0\}\) is arbitrary, this completes the proof. \(\square\)

**Lemma 4.8** For any \((\hat{v}, \hat{q}), (\hat{w}, \hat{r}) \in W_{k+d}\), we have
\[ a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r})) \leq c_3 \| (\hat{w}, \hat{r}) \|_D \| (\hat{v}, \hat{q}) \|_D, \tag{4.31} \]
where \(c_3\) is a positive constant independent of \(h\).

**Proof** We continue to use \(\{\varphi_j\}^{N_v}_{j=1}\) and \(\{\psi_j\}^{N_p}_{j=1}\) as the bases in space \(W_{k+d}\) for velocity and pressure, respectively.

Let
\[
\hat{v} = \sum_{j=1}^{N_v} x_{v,j} \varphi_j, \quad x_v = (x_{v,1}, x_{v,2}, \ldots, x_{v,N_v})^T, \\
\hat{q} = \sum_{j=1}^{N_p} x_{p,j} \psi_j, \quad x_p = (x_{p,1}, x_{p,2}, \ldots, x_{p,N_p})^T, \\
\hat{w} = \sum_{j=1}^{N_v} y_{v,j} \varphi_j, \quad y_v = (y_{v,1}, y_{v,2}, \ldots, y_{v,N_v})^T, \\
\hat{r} = \sum_{j=1}^{N_p} y_{p,j} \psi_j, \quad y_p = (y_{p,1}, y_{p,2}, \ldots, y_{p,N_p})^T.
\]

Then, using (4.24), Cauchy-Schwarz inequality, Lemmas 4.5, 4.1, and 4.6
\[
a_3((\hat{v}, \hat{q}), (\hat{w}, \hat{r})) = y_v^T D_v x_v + y_p^T B x_p - y_p^T B T x_v + y_p^T (c_s D_p - B^T D_v^{-1} B) x_p \\
\leq |\hat{v}|_D |\hat{w}|_D + \|q\| \|\nabla \cdot \hat{w}\| + \|\hat{r}\| \|\nabla \cdot \hat{v}\| + c_s \|\hat{q}\|_p \|\hat{r}\|_p + \|\hat{q}\|_B \|\hat{r}\|_B \\
\leq |\hat{v}|_D |\hat{w}|_D + d \|\hat{q}\| \|\hat{w}\|_{1,\Omega} + d \|\hat{r}\| \|\hat{v}\|_{1,\Omega} + (c_s c_i^{-1} + 1) \|\hat{q}\|_B \|\hat{r}\|_B \\
\leq |\hat{v}|_D |\hat{w}|_D + d \beta_2 \|\hat{q}\| \|\hat{w}\|_D + d \beta_2 \|\hat{r}\| \|\hat{v}\|_D + (c_s c_i^{-1} + 1) d(d + 1) \beta_2 \|\hat{q}\| \|\hat{r}\| \\
\leq c_3 \| (\hat{w}, \hat{r}) \|_D \| (\hat{v}, \hat{q}) \|_D, \tag{4.42}
\]
with \(c_3 = \sqrt{2} \max \{1, d \beta_2, (c_s c_i^{-1} + 1) d(d + 1) \beta_2\}\). \(\square\)

**Lemma 4.9** Let \((\bar{e}_u, \bar{e}_p)\) and \((\tilde{e}_u, \tilde{e}_p)\) be the solutions of (4.8) and (4.22)-(4.23), respectively. Then,
\[
\frac{(\mu \beta_1)^2}{c_2(1 + \mu \beta_1)^2} \| (\bar{e}_u, \bar{e}_p) \|_D \leq \| (\bar{e}_u, \bar{e}_p) \|_D \leq \frac{c_3(1 + \mu \beta_1)^2}{(\mu \beta_1)^2} \| (\bar{e}_u, \bar{e}_p) \|_D, \tag{4.32}
\]
where \(\| \cdot \|_D\) is defined in (4.6). The constants \(c_2, c_3, \beta_1, \) and \(\mu\) are defined in (4.13), (4.31), (4.11), and (2.16), respectively.
Proof It follows from (4.8) and (4.25) that

$$a_3((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q})) = a_2((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q})), \quad \forall (\tilde{v}, \tilde{q}) \in W_{k+d}. \quad (4.33)$$

Using (4.33) and Lemma 4.7, we obtain

$$\frac{(\mu \beta_1)^2}{(1 + \mu \beta_1)^2} \| (\tilde{e}_u, \tilde{e}_p) \|_D \leq \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \frac{a_3((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q}))}{\| (\tilde{v}, \tilde{q}) \|_D}$$

$$= \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \frac{a_2((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q}))}{\| (\tilde{v}, \tilde{q}) \|_D}$$

$$\leq \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \mathcal{C}_2 \| (\tilde{e}_u, \tilde{e}_p) \|_D \| (\tilde{v}, \tilde{q}) \|_D$$

which implies the first inequality in (4.32).

Similarly, using (4.33) and Lemma 4.3, we have

$$\frac{(\mu \beta_1)^2}{(1 + \mu \beta_1)^2} \| (\tilde{e}_u, \tilde{e}_p) \|_D \leq \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \frac{a_3((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q}))}{\| (\tilde{v}, \tilde{q}) \|_D}$$

$$= \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \frac{a_2((\tilde{e}_u, \tilde{e}_p), (\tilde{v}, \tilde{q}))}{\| (\tilde{v}, \tilde{q}) \|_D}$$

$$\leq \sup_{(\tilde{v}, \tilde{q}) \in W_{k+d} \setminus \{0\}} \mathcal{C}_3 \| (\tilde{e}_u, \tilde{e}_p) \|_D \| (\tilde{v}, \tilde{q}) \|_D$$

which implies the second inequality in (4.32).

Combining Theorem 4.2 and Lemma 4.9, we obtain the following lower and upper bounds for the estimator \(\|(\tilde{e}_u, \tilde{e}_p)\|_D\).

Theorem 4.4 Let \((u, p), (\hat{u}, \hat{p})\) and \((\hat{e}_u, \hat{e}_p)\) be the solutions of (2.1), (2.15), and (4.25), respectively. There are constants \(\mathcal{E}_*\) and \(\bar{\mathcal{E}}^*\) such that

\[
\mathcal{E}_* \| (\hat{e}_u, \hat{e}_p) \|_D \leq \| (u - \hat{u}, p - \hat{p}) \| \leq \bar{\mathcal{E}}^* \| (\hat{e}_u, \hat{e}_p) \|_D + \frac{C_T \text{osc}(f)}{\epsilon_1},
\]

where \(\| \cdot \|_V\) and \(\| \cdot \|_D\) are defined in (2.2) and (4.6), respectively. The constants \(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{C}_1, \mu, \beta_1, \beta_2,\) and \(C_T\) are defined in (2.3), (4.13), (4.31), (2.4), (2.16), (4.11), and Lemma 3.1, respectively.

5 Adaptive Algorithm

In this section, we construct an adaptive FEM to solve (1.1)-(1.3) based on the local and global \textit{a posteriori} error estimators, denoted by \(\eta_L, T\) and \(\eta_G(T_m)\), defined in (5.1) and (5.2), which produce a sequence of discrete solutions \((\hat{u}_m, \hat{p}_m)\) in nested spaces \(V_{k,m}\) over triangulation
The index $m$ indicates the underlying mesh with size $h_m$. Assume that an initial mesh $T_0$, a Döfler parameter $\theta \in (0, 1)$, and a targeted tolerance $\varepsilon$ are given.

Actually, a common adaptive refinement scheme involves a loop structure of the form:

\[
\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}
\]

with the initial triangulation $T_0$ of $\Omega$ (cf. [18, 20]). \text{SOLVE} refers to solving the FEM scheme (2.8) on a relatively coarse mesh $T_m$. \text{ESTIMATE} relies on an efficient and reliable a posteriori error estimate, and the local and global estimators are defined in (5.1) and (5.2). With the help of the error estimators, \text{MARK} determines the elements to be refined, hence creating a subset $S_m$ of $T_m$ for refinement. Finally, \text{REFINE} generates a finer triangulation $T_{m+1}$ by dividing those elements in $S_m$, and an updated numerical solution will be computed on $T_{m+1}$.

For the first and the last step, there have been rapid advances for solving the linear system (2.8) and refinement implementation, respectively in recent years, and we refer to [8, 9, 26] for the details. Here, we focus on the interplay between the error estimator and the marking strategy. The error estimator consists of local and global estimators for a given triangulation. The local estimator provides the information for the marking strategy to determine the triangles to be refined, while the global error estimator provides the measure for the reliable stop condition of the loops.

Recall that for the three error problems (2.15), (4.8), and (4.25) with the associated norm (2.2) and (4.6), we define the local and global error estimator with three different cases as

\[
\eta_{L,T} = \begin{cases} 
\sqrt{\|\hat{e}_u\|_{1,T}^2 + \|\hat{e}_p\|_{0,T}^2}, & \text{Case 1,} \\
\sqrt{\|\hat{e}_u\|_{2,D,T}^2 + \|\hat{e}_p\|_{0,T}^2}, & \text{Case 2,} \\
\sqrt{\|\hat{e}_u\|_{2,D,T}^2 + \|\tilde{e}_p\|_{0,T}^2}, & \text{Case 3,}
\end{cases}
\]  

(5.1)

and

\[
\eta_G(T_m) = \sqrt{\sum_{T \in T_m} \eta_{L,T}^2}.
\]  

(5.2)

By definition (2.2) and (4.6), in Case 1, Case 2, and Case 3, $\eta_G(T_m)$ will be $\|(\hat{e}_u, \hat{e}_p)\|_V$, $\|(\hat{e}_u, \hat{e}_p)\|_D$, and $\|(\hat{e}_u, \hat{e}_p)\|_D$, respectively. Based on Theorems 3.1, 4.2, and 4.4, the global error estimator $\eta_G(T_m)$ provides an estimate of the discretization error $\|(u - \hat{u}, p - \hat{p})\|_V$, which is frequently used to judge the quality of the underlying discretization. The local error estimator $\eta_{L,T}$ is an estimate of the error on element $T$. All elements $T \in T_m$ are marked for refinement, if $\eta_{L,T}$ exceeds the certain tolerance. Denote the set of all marked elements by $S_m \subset T_m$. The global error estimator associated with $S_m$ is denoted by $\eta_G(S_m)$.

The algorithm of adaptive FEM is listed in Algorithm 1.

In the \text{MARK} step, we adopt the Döfler marking strategy which is a mature strategy and is widely used in the adaptive algorithm [12]. Recently, it has been shown that Döfler marking with minimal cardinality is a linear complexity problem [21]. In this marking strategy the local error estimators $\{\eta_{L,T}\}_{T \in T_m}$ are sorted in descending order. The sorted local error estimators are denoted by $\{\tilde{\eta}_{L,T}\}_{T \in T_m}$. Then, the set of elements marked for refinement is given by $\{\tilde{\eta}_{L,T}\}_{T \in S_m}$, where $S_m$ contains the least number of elements such that

\[
\eta_G^2(S_m) = \sum_{T \in S_m} \tilde{\eta}_{L,T}^2 \geq \theta \eta_G^2(T_m).
\]
Algorithm 1 Adaptive FEM

Input: Construct an initial mesh $T_0$. Choose a parameter $0 < \theta < 1$ and a tolerance $\varepsilon$.

Output: Final triangulation $T_M$ and the finite element approximation $(\hat{u}_M, \hat{p}_M)$ on $T_M$.

Set $m = 0$ and $\eta_G(T_m) = 1$.

While $\eta_G(T_m) > \varepsilon$

1. (SOLVE) Solve the FEM scheme (2.8) on $T_m$.
2. (ESTIMATE) Compute the local error estimator as defined in (5.1) for all elements $T \in T_m$.
3. (MARK) Construct a subset $S_m \subset T_m$ with least number of elements such that

$$\eta^2_G(S_m) \geq \theta \eta^2_G(T_m).$$

4. (REFINE) Refine elements in $S_m$ together with the elements, which must be refined to make $T_{m+1}$ conforming.
5. Set $m = m + 1$.

End

Set $(\hat{u}_M, \hat{p}_M) = (\hat{u}_m, \hat{p}_m)$ and $T_M = T_m$.

Generally speaking, a small value of $\theta$ leads to a small set $S_m$, while a large value of $\theta$ leads to a large set $S_m$. In [12], $\theta$ is suggested to be adopted in $[0.5, 0.8]$. We emphasize that many auxiliary elements are refined to eliminate the hanging nodes, which may have been created in the MARK step. There are many mature toolkits to process the hanging nodes [26].

Finally, to show the effectiveness of the global error estimator defined in (5.2), we introduce the effective index as follows

$$\kappa_{eff} = \frac{\eta_G(T_m)}{\|(u - \hat{u}, p - \hat{p})\|_V},$$

which is the ratio between the global error estimator and the FEM approximation error. According to Theorems 3.1, 4.2, and 4.4, the effective index is bounded from both above and below.

6 Numerical Experiments

In this section, we present two-dimensional numerical examples to verify the properties of the estimators mentioned in Theorems 3.1, 4.2, and 4.4 using uniform and adaptive refinement. All these simulations have been implemented on a 3.2GHz quad-core processor with 16GB RAM by Matlab.

Example 1. Let $\Omega$ be the square domain $(-1, 1) \times (-1, 1)$. Further, select $f = 0$ and enforce appropriate inhomogeneous boundary conditions for $\mathbf{u}$ on $\Gamma$ so that the analytical solution to (1.1)-(1.3) is given by

\[ u_1(x, y) = -e^x(y \cos(y) + \sin(y)), \]
\[ u_2(x, y) = e^x y \sin(y), \]
\[ p(x, y) = 2e^x \sin(y). \]

In this example, we test three different combinations of finite element spaces. More specifically, the finite element space with respect to velocity and pressure for approximation and error problems are $(V_1, W_3)$, $(V_2, W_4)$, and $(V_3, W_5)$, respectively, which are defined in (2.7) and (2.14).
Figure 1 shows the bases of \((WV_4, WP_3)\), \((WV_5, WP_4)\), and \((WV_6, WP_5)\) defined in (2.12) and (2.13) from left to right, respectively. The Lagrange bases for Taylor-Hood element \((VV_2, VP_1)\), \((VV_3, VP_2)\), and \((VV_4, VP_3)\) defined in (2.5) and (2.6) can also be found in [19]. Table 1 shows the degree of freedom \((d.o.f)\) of any component with respect to velocity and pressure in any element. In Table 2, the error \(\| (u - \hat{u}, p - \hat{p}) \| \) is compared with three \textit{a posteriori} error estimators from (2.15), (4.8), and (4.25), respectively. We can conclude that the error \(\| (u - \hat{u}, p - \hat{p}) \| \) has the same convergence order as our proposed three error estimators with respect to \(d.o.f\), and the convergence order is optimal. These are consistent with Theorems 3.1, 4.2, and 4.4.

The effective index and computational cost of the three \textit{a posteriori} error estimators are shown in Table 3. In each of cases of (5.1), the effective index converges to a constant, which is case dependent. The computational cost of the first error estimator is comparable to the approximation problem, while that of the latter two error estimators is very small, relative to the approximation problem. In particular, the computational cost of the third error estimator increases slowly with the increase of \(d.o.f\). The reason is that we need to solve only two global diagonal linear systems corresponding to the degree of freedom of velocity and pressure, respectively, to get the third error estimator. In conclusion, these results fully confirm that the diagonalization technologies of velocity and pressure introduced in Sect. 4 are effective and successful. A more intuitive display of convergence and effective index can be found in Fig. 2.

\textbf{Example 2.} This example is taken from page 113 in [25]. The solution is singular at the origin. Let \(\Omega\) be the L-shape domain \((-1, 1)^2 \setminus [0, 1) \times (-1, 0]\), and select \(f = 0\). Then, use \((r, \varphi)\) to denote the polar coordinates. We impose an appropriate inhomogeneous boundary conditions...
The errors, the global error estimators, and the convergence orders in Example 1

| d.o.f       | $\|u - \hat{u}, - \hat{p}\|_V$ | Order | $\|\hat{e}_u, \hat{e}_p\|_V$ | Order | $\|\hat{e}_u, \hat{e}_p\|_D$ | Order | $\|\hat{e}_u, \hat{e}_p\|_D$ | Order |
|-------------|---------------------------------|-------|-----------------------------|-------|-----------------------------|-------|-----------------------------|-------|
| $V_1-W_3$   |                                 |       |                             |       |                             |       |                             |       |
| 95          | 4.188E-1                        | 4.928E-1 | 3.32E-1                    |       | 2.318E-1                    |       |                             |       |
| 331         | 6.578E-2                        | 1.115E-1 | 7.295E-2                   | 1.217 | 6.507E-2                    | 1.017 |                             |       |
| 1234        | 1.213E-2                        | 2.846E-2 | 1.796E-2                   | 1.064 | 1.724E-2                    | 1.008 |                             |       |
| 4771        | 2.619E-3                        | 7.311E-3 | 4.523E-3                   | 1.020 | 4.445E-3                    | 1.003 |                             |       |
| 18755       | 6.169E-4                        | 1.858E-3 | 1.137E-3                   | 1.008 | 1.238E-3                    | 1.001 |                             |       |
| 74371       | 1.505E-4                        | 4.689E-4 | 2.855E-4                   | 1.003 | 2.844E-4                    | 1.000 |                             |       |
| $V_2-W_4$   |                                 |       |                             |       |                             |       |                             |       |
| 211         | 3.101E-2                        | 5.517E-2 | 4.928E-2                   |       | 2.779E-2                    |       |                             |       |
| 771         | 4.043E-3                        | 8.217E-3 | 6.305E-3                   | 1.586 | 3.593E-3                    | 1.578 |                             |       |
| 2947        | 4.794E-4                        | 1.105E-3 | 8.148E-4                   | 1.526 | 4.482E-4                    | 1.552 |                             |       |
| 11523       | 45.684E-5                       | 1.425E-4 | 1.034E-4                   | 1.513 | 5.546E-5                    | 1.532 |                             |       |
| 45571       | 6.865E-6                        | 1.806E-5 | 1.302E-5                   | 1.507 | 6.877E-6                    | 1.518 |                             |       |
| 181251      | 8.415E-7                        | 2.727E-6 | 1.632E-5                   | 1.504 | 8.554E-7                    | 1.509 |                             |       |
| $V_3-W_5$   |                                 |       |                             |       |                             |       |                             |       |
| 375         | 3.698E-3                        | 4.145E-3 | 4.420E-3                   |       | 1.513E-3                    |       |                             |       |
| 1403        | 1.598E-4                        | 2.408E-4 | 2.886E-4                   | 2.068 | 8.777E-5                    | 2.158 |                             |       |
| 5427        | 8.997E-6                        | 1.519E-5 | 1.871E-5                   | 2.022 | 5.427E-6                    | 2.057 |                             |       |
| 21347       | 5.521E-7                        | 9.643E-7 | 1.206E-6                   | 2.001 | 3.398E-7                    | 2.023 |                             |       |
| 84675       | 3.447E-8                        | 6.077E-8 | 7.647E-8                   | 2.002 | 2.128E-8                    | 2.010 |                             |       |
| 337283      | 2.171E-9                        | 3.814E-9 | 4.808E-9                   | 2.001 | 1.331E-9                    | 2.005 |                             |       |
Table 3  The effective index and computational cost in Example 1

| d.o.f  | Effective index | Time(s) | Problem (2.8) | Problem (2.15) | Problem (4.8) | Problem (4.25) |
|--------|-----------------|---------|---------------|---------------|---------------|---------------|
|        | Case 1 | Case 2 | Case 3 | Case 1 | Case 2 | Case 3 | Case 1 | Case 2 | Case 3 | Case 1 | Case 2 | Case 3 | Case 1 | Case 2 | Case 3 |
| $V_1-W_3$ | 95     | 1.176  | 0.795  | 0.553  | 0.041  | 0.024  | 0.006  | 0.007  | 331     | 1.695  | 1.108  | 0.989  | 0.040  | 0.028  | 0.006  | 0.007  | 1234    | 2.345  | 1.481  | 1.421  | 0.055  | 0.050  | 0.008  | 0.007  | 4771    | 2.791  | 1.726  | 1.697  | 0.138  | 0.178  | 0.023  | 0.022  | 18755   | 3.013  | 1.844  | 1.829  | 0.551  | 0.641  | 0.025  | 0.015  | 74371   | 3.114  | 1.895  | 1.888  | 3.313  | 3.254  | 0.097  | 0.047  |
| $V_2-W_4$ | 211     | 1.779  | 1.589  | 0.896  | 0.089  | 0.044  | 0.015  | 0.016  | 771     | 2.032  | 1.559  | 0.888  | 0.092  | 0.057  | 0.016  | 0.020  | 2947    | 2.305  | 1.699  | 0.934  | 0.147  | 0.095  | 0.020  | 0.019  | 11523   | 2.508  | 1.820  | 0.975  | 0.435  | 0.311  | 0.043  | 0.034  | 45571   | 2.631  | 1.896  | 1.001  | 1.997  | 1.377  | 0.098  | 0.044  | 181251  | 2.699  | 1.939  | 1.016  | 15.441 | 8.122  | 0.474  | 0.127  |
| $V_3-W_5$ | 375     | 1.120  | 1.119  | 0.409  | 0.126  | 0.051  | 0.021  | 0.020  | 1403    | 1.507  | 1.805  | 0.549  | 0.144  | 0.065  | 0.023  | 0.022  | 5427    | 1.688  | 2.079  | 0.603  | 0.246  | 0.139  | 0.031  | 0.024  | 21347   | 1.746  | 2.185  | 0.615  | 0.859  | 0.452  | 0.074  | 0.047  | 84675   | 1.762  | 2.218  | 0.617  | 3.734  | 1.830  | 0.212  | 0.072  | 337283  | 1.756  | 2.214  | 0.613  | 23.286 | 10.231 | 1.058  | 0.266  |
condition for $u$ so that

$$u_1(r, \varphi) = r^\lambda ((1 + \lambda) \sin(\varphi)\Psi(\varphi) + \cos(\varphi)\Psi'(\varphi)),$$

$$u_2(r, \varphi) = r^\lambda (\sin(\varphi)\Psi'(\varphi) - (1 + \lambda) \cos(\varphi)\Psi(\varphi)),$$

$$p(r, \varphi) = -r^{\lambda - 1} [(1 + \lambda)^2 \Psi'(\varphi) + \Psi'']/(1 - \lambda),$$

where

$$\Psi(\varphi) = \sin((1 + \lambda)\varphi) \cos(\lambda \omega)/(1 + \lambda) - \cos((1 + \lambda)\varphi)$$

$$- \sin((1 - \lambda)\varphi) \cos(\lambda \omega)/(1 - \lambda) + \cos((1 - \lambda)\varphi),$$

$$\omega = \frac{3\pi}{2}.$$

The exponent $\lambda$ is the smallest positive solution of

$$\sin(\lambda \omega) + \lambda \sin(\omega) = 0,$$

thereby, $\lambda \approx 0.54448373678246$.

We emphasize that $(u, p)$ is analytic in $\overline{\Omega}\setminus\{0\}$, but both $\nabla u$ and $p$ are singular at the origin; indeed, here $u \notin \{H^2(\Omega)\}^2$ and $p \notin H^1(\Omega)$. This example reflects the typical (singular) behavior that solutions of the two-dimensional Stokes problem exhibit in the vicinity of reentrant corners in the computational domain.

We denote the finite element spaces by $V_1$ and $W_3$ in the approximation problem and the error problem, respectively. The finite element space $V_1$ consists of velocity space and pressure space. The velocity space is the space of continuous piecewise quadratic polynomials and the pressure space is the space of continuous piecewise linear polynomials associated with $T$. In this example, we will test three different cases in (5.1). In the MARK step, we denote the Döfler parameter by $\theta$. In the REFINE step, the refinement process is implemented using the MATLAB function REFINEMESH. The key is dividing the marked element into four parts by regular refinement (dividing all edges of the selected triangles in half).
Figure 3 shows that for Example 2, $\eta_G(T_m)$ has close convergent rates near $-1$ with respect to the d.o.f for different $\theta$ in three cases in (5.1). However, the convergence order of the uniform refinement is only $-0.28$, as can be seen from Fig. 4. From the comparison of computational cost for different $\theta$ when $\|(u - \hat{u}, p - \hat{p})\|_V < 0.25$ in Table 4, adaptive FEM is much faster than the uniform refinement. From Tables 4 and 5, we can conclude that the three a posteriori error estimators have similar error performance and similar computational cost for approximation problem (2.8) in each step, but there is a big gap in computational cost for error problems. Computational cost for the latter two cases are much cheaper than the first one. Especially, for large d.o.f, computational cost for the third case is much cheaper than the first two cases.

Next, we set $\theta = 0.5$ and compare the convergent rates between $\|(u - \hat{u}, p - \hat{p})\|_V$ and $\eta_G(T)$ for three cases shown in Fig. 4. The x-axes denotes the d.o.f in log scale, while y-axes denotes the errors in log scale. The squared line denotes the error $\|(u - \hat{u}, p - \hat{p})\|_V$ of the uniform refinement. The pentagram, asterisk, and the circled lines denote error $\|(u - \hat{u}, p - \hat{p})\|_V$ and $\eta_G(T_m)$ of adaptive FEM for the three cases, respectively. It is obvious that $\|(u - \hat{u}, p - \hat{p})\|_V$ and $\eta_G(T_m)$ have the same convergence order and have a higher convergence order than the uniform refinement.

For Case 3, Fig. 5 shows the initial mesh with d.o.f = 259, eighth refinement mesh with d.o.f = 2465, and fourteenth refinement mesh with d.o.f = 19950. It has inserted refinement elements around the singularity at $(x, y) = (0, 0)$ as d.o.f increases to reduce the global error. Tables 6, 7, and 8 show the error $\|(u - \hat{u}, p - \hat{p})\|_V$, the global error estimator...
Table 4 Computational cost for different $\theta$ in Example 2

| $\theta$ | Refinement steps | $d.o.f$ | $\| (u - \hat{u}, p - \hat{p}) \|_V$ | Time(s) of (2.8) | Time(s) of $\eta_G(T)$ |
|----------|------------------|--------|---------------------------------|-----------------|-------------------|
| Case 1   | 0.9 7            | 3216   | 0.245                           | 0.733           | 0.923             |
|          | 0.8 7            | 2629   | 0.243                           | 0.698           | 0.872             |
|          | 0.7 8            | 2756   | 0.244                           | 0.763           | 0.944             |
|          | 0.6 8            | 2238   | 0.181                           | 0.736           | 0.893             |
|          | 0.5 8            | 1772   | 0.185                           | 0.753           | 0.859             |
| Case 2   | 0.9 7            | 3220   | 0.246                           | 0.771           | 0.166             |
|          | 0.8 7            | 2476   | 0.249                           | 0.634           | 0.169             |
|          | 0.7 8            | 2896   | 0.173                           | 0.722           | 0.171             |
|          | 0.6 8            | 2138   | 0.181                           | 0.715           | 0.174             |
|          | 0.5 8            | 1772   | 0.186                           | 0.709           | 0.178             |
| Case 3   | 0.9 7            | 3230   | 0.246                           | 0.692           | 0.168             |
|          | 0.8 8            | 3500   | 0.171                           | 0.800           | 0.169             |
|          | 0.7 8            | 2729   | 0.174                           | 0.754           | 0.165             |
|          | 0.6 8            | 2120   | 0.180                           | 0.739           | 0.161             |
|          | 0.5 8            | 1758   | 0.187                           | 0.708           | 0.163             |
| uniform  | 1.0 7            | 887299 | 0.243                           | 130.463         | —                 |

Table 5 Computational cost for different $a$ posteriori estimators with $\theta = 0.5$ in Example 2

| Refinement steps | $d.o.f$ | $\| (u - \hat{u}, p - \hat{p}) \|_V$ | $\eta_G(T_m)$ | Time(s) of (2.8) | Time(s) of $\eta_G(T_m)$ |
|------------------|--------|---------------------------------|---------------|-----------------|-------------------|
| Case 1           | 19     | 99956 2.976E-3                   | 2.845E-3      | 19.868          | 22.238            |
| Case 2           | 19     | 94765 3.035E-3                   | 2.238E-3      | 16.702          | 1.227             |
| Case 3           | 19     | 91087 3.081E-3                   | 2.521E-3      | 14.250          | 0.512             |

Fig. 5 The meshes with $d.o.f = 259$ (left), 2462 (middle), and 19950 (right) for Example 2

$\eta_G(T_m)$, and the effective index of the adaptive FEM as the $d.o.f$ increases for three cases in (5.1). The results of effective index $\kappa_{eff}$ defined in (5.3) are shown in the sixth column. It is bounded with adaptive refinement, which shows that the estimator is reliable.

Example 3. In this case, we test the lid-driven cavity problem. The domain is taken as the square $\Omega = (0, 1) \times (0, 1)$, we set $f = 0$, and the boundary conditions $u = 0$ on $[[0] \times (0, 1)] \cup [(0, 1) \times [0]] \cup [(1] \times (0, 1)]$ and $u = (1, 0)^T$ on $(0, 1) \times \{1\}$. This problem
The errors, the global error estimators, and the effective index for Case 1 in Example 2

| d.o.f | $(\|u - \hat{u}, p - \hat{p}\|_V)$ | Order | $(\|\hat{e}_u, \hat{e}_p\|_V)$ | Order | $\kappa_{eff}$ |
|-------|----------------------------------|--------|---------------------------------|--------|----------------|
| 259   | 2.452E0                          | —      | 1.540E0                         | —      | 0.628          |
| 426   | 1.644E0                          | 0.803  | 1.093E0                         | 0.689  | 0.664          |
| 593   | 1.116E0                          | 1.171  | 7.680E-1                        | 1.068  | 0.687          |
| 760   | 7.640E-1                         | 1.528  | 5.404E-1                        | 1.416  | 0.707          |
| 927   | 5.248E-1                         | 1.890  | 3.866E-1                        | 1.686  | 0.736          |
| 1094  | 3.686E-1                         | 2.132  | 2.861E-1                        | 1.816  | 0.776          |
| 1261  | 2.627E-1                         | 2.385  | 2.236E-1                        | 1.736  | 0.851          |
| 1772  | 1.850E-1                         | 1.030  | 1.607E-1                        | 0.970  | 0.868          |
| 2597  | 1.264E-1                         | 0.996  | 1.084E-1                        | 1.029  | 0.857          |
| 3401  | 8.767E-2                         | 1.357  | 7.868E-2                        | 1.189  | 0.897          |
| 4933  | 6.030E-2                         | 1.006  | 5.550E-2                        | 0.938  | 0.920          |
| 7503  | 4.121E-2                         | 0.907  | 3.777E-2                        | 0.917  | 0.916          |
| 10687 | 2.897E-2                         | 0.996  | 2.684E-2                        | 0.965  | 0.926          |
| 15487 | 1.959E-2                         | 1.054  | 1.827E-2                        | 1.037  | 0.932          |
| 22647 | 1.357E-2                         | 0.966  | 1.263E-2                        | 0.970  | 0.930          |
| 33282 | 9.305E-3                         | 0.980  | 8.809E-3                        | 0.936  | 0.946          |
| 47811 | 6.413E-3                         | 1.027  | 6.048E-3                        | 1.038  | 0.943          |
| 67688 | 4.355E-3                         | 1.112  | 4.135E-3                        | 1.093  | 0.949          |
| 99956 | 2.976E-3                         | 0.976  | 2.845E-3                        | 0.959  | 0.956          |

The errors, the global error estimators, and the effective index for Case 2 in Example 2

| d.o.f | $(\|u - \hat{u}, p - \hat{p}\|_V)$ | order | $(\|\hat{e}_u, \hat{e}_p\|_D)$ | order | $\kappa_{eff}$ |
|-------|----------------------------------|--------|---------------------------------|--------|----------------|
| 259   | 2.452E0                          | —      | 1.296E0                         | —      | 0.528          |
| 426   | 1.644E0                          | 0.803  | 9.179E-1                        | 0.694  | 0.558          |
| 593   | 1.116E0                          | 1.171  | 6.426E-1                        | 1.077  | 0.575          |
| 760   | 7.640E-1                         | 1.528  | 4.496E-1                        | 1.439  | 0.588          |
| 927   | 5.248E-1                         | 1.890  | 3.179E-1                        | 1.743  | 0.605          |
| 1094  | 3.686E-1                         | 2.132  | 2.307E-1                        | 1.936  | 0.625          |
| 1261  | 2.627E-1                         | 2.385  | 1.750E-1                        | 1.944  | 0.666          |
| 1772  | 1.865E-1                         | 1.006  | 1.257E-1                        | 0.971  | 0.674          |
| 2458  | 1.281E-1                         | 1.148  | 8.717E-2                        | 1.119  | 0.680          |
| 3305  | 8.803E-2                         | 1.267  | 6.103E-2                        | 1.203  | 0.693          |
| 4609  | 6.158E-2                         | 1.074  | 4.348E-2                        | 1.01   | 0.706          |
| 7001  | 4.176E-2                         | 0.928  | 2.999E-2                        | 0.888  | 0.718          |
| 10508 | 2.912E-2                         | 0.888  | 2.081E-2                        | 0.899  | 0.714          |
| 15163 | 1.979E-2                         | 1.053  | 1.413E-2                        | 1.054  | 0.714          |
| 22102 | 1.355E-2                         | 1.003  | 9.819E-3                        | 0.967  | 0.724          |
| 32394 | 9.325E-3                         | 0.979  | 6.748E-3                        | 0.981  | 0.723          |
| 46481 | 6.415E-3                         | 1.035  | 4.652E-3                        | 1.029  | 0.725          |
| 64824 | 4.430E-3                         | 1.113  | 3.255E-3                        | 1.074  | 0.734          |
| 94765 | 3.035E-3                         | 0.995  | 2.238E-3                        | 0.985  | 0.737          |
Table 8 The errors, the global error estimators, and the effective index for Case 3 in Example 2

| d.o.f | $\| (\mathbf{u} - \hat{\mathbf{u}}, p - \hat{p}) \|_V$ | Order | $\| (\mathbf{e}_u, \mathbf{e}_p) \|_D$ | Order | $\kappa_{eff}$ |
|-------|---------------------------------|-------|---------------------------------|-------|-----------|
| 259   | 2.452E0                         | —     | 1.463E0                         | —     | 0.596     |
| 426   | 1.644E0                         | 0.803 | 1.024E0                         | 0.715 | 0.622     |
| 593   | 1.116E0                         | 1.171 | 7.169E-1                        | 1.079 | 0.642     |
| 760   | 7.640E-1                        | 1.528 | 5.011E-1                        | 1.443 | 0.655     |
| 927   | 5.248E-1                        | 1.890 | 3.537E-1                        | 1.752 | 0.674     |
| 1094  | 3.686E-1                        | 2.132 | 2.558E-1                        | 1.956 | 0.693     |
| 1261  | 2.627E-1                        | 2.385 | 1.930E-1                        | 1.981 | 0.734     |
| 1758  | 1.871E-1                        | 1.020 | 1.394E-1                        | 0.978 | 0.745     |
| 2462  | 1.301E-1                        | 1.078 | 9.810E-2                        | 1.045 | 0.753     |
| 3477  | 8.875E-2                        | 1.109 | 6.862E-2                        | 1.035 | 0.773     |
| 4573  | 6.154E-2                        | 1.336 | 4.936E-2                        | 1.201 | 0.802     |
| 6511  | 4.357E-2                        | 0.976 | 3.466E-2                        | 1.001 | 0.795     |
| 9702  | 3.005E-2                        | 0.931 | 2.369E-2                        | 0.953 | 0.788     |
| 14095 | 2.027E-2                        | 1.053 | 1.642E-2                        | 0.981 | 0.809     |
| 19950 | 1.396E-2                        | 1.073 | 1.136E-2                        | 1.060 | 0.813     |
| 29128 | 9.611E-3                        | 0.986 | 7.781E-3                        | 1.000 | 0.809     |
| 42657 | 6.514E-3                        | 1.019 | 5.336E-3                        | 0.988 | 0.819     |
| 62401 | 4.477E-3                        | 0.986 | 3.650E-3                        | 0.997 | 0.815     |
| 91087 | 3.081E-3                        | 0.988 | 2.521E-3                        | 0.978 | 0.818     |

Fig. 6  The initial mesh and tenth refinement mesh for Example 3

has a corner singularity. The tangential component of velocity $\mathbf{u} \cdot \mathbf{\tau}$ has a discontinuity at the two top corners, where $\mathbf{\tau}$ denotes the unit tangential vector on the boundary. We use the adaptive FEM algorithm with the third estimator to solve this problem. The finite element space and refinement criterion are the same as in Example 2. The Döfler parameter is $\theta = 0.5$. Figure 6 shows that the refinement of mesh focuses on the two top corners. In Fig. 7, we depict the discrete pressure field obtained using the initial and adapted meshes where we note the improvement in the quality of the computed solution since the singular nature of the pressure is better captured in the adapted mesh.
7 Conclusion

In this paper, we present \textit{a posteriori} error estimator for the Stokes problem with Dirichlet boundary condition. Based on auxiliary subspace techniques, we proposed a hierarchical basis \textit{a posteriori} error estimator, which is proved to have global upper and lower bounds without saturation assumption. In order to reduce the computational cost, we diagonalize the error problem with respect to velocity and pressure in sequence. The diagonalized \textit{a posteriori} error estimator needs to solve only two global diagonal linear systems. And it also has global upper and lower bounds. Numerical experiments are shown to illustrate the reliability of the \textit{a posteriori} error estimator using uniform and adaptive refinement.

Acknowledgements The work of J.C. Zhang was supported by the Natural Science Foundation of Jiangsu Province (grant no.BK20210540) and The Natural Science Foundation of the Jiangsu Higher Education Institutions of China (grant no.21KJB110015). The work of R. Zhang was supported by the National Key Research and Development Program of China (grant no.2020YFA0713601).

Funding The authors have not disclosed any funding.

Data Availability Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

Appendix A.

The proof of Lemma 2.2 The idea of proof is similar to [8] for $d = 2$ and [9] for $d = 3$. Next, we will give proofs for $d = 2$ and $d = 3$, respectively.

Case $d = 2$: The idea is to consider a macroelement partition of the domain $\Omega$ in such a way that each macroelement contains exactly three triangles. By virtue of Remark 3.3 in [8], it suffices to prove the inf-sup condition for only one macroelement. We consider a macroelement $\Omega_i = a \cup b \cup c$ as in Fig. 8.

Let us introduce some notations. We denote by $\lambda^d_{AB}$ the barycentric coordinate related to the triangle $a$, which vanishes on the edge $AB$ (analogous notations for the other cases).
Fig. 8 The macroelement partition containing three triangles

we denote by $L_{i,x}^a$ the $i$-th Legendre polynomial with respect to the measure $\mu_{a,x}$ defined as

$$\int_{x_A}^0 f(x)d\mu_{a,x} = \int_a^a \lambda_{AB}^a \lambda_{AE}^a f(x)dxdy \ \forall f(x) : [x_A, 0] \rightarrow \mathbb{R}, \quad (A.1)$$

where $x_A$ is the x-coordinate of the vertex $A$. A similar definition will hold for $L_{i,y}^c$ using $\lambda_{BC}^c \lambda_{CD}^c$. On the triangle $b$ we shall use both $L_{i,y}^b$ (using $\lambda_{ED}^b \lambda_{BD}^b$) and $L_{i,y}^b$ (using $\lambda_{BE}^b \lambda_{ED}^b$). These Legendre polynomials are defined up to a constant factor so that we can normalize them by imposing that they assume the same value at the origin. This is possible by virtue of Proposition 2.1 in [8].

Our approach to the stability condition will be related to the modified inf-sup condition that can be written as

$$\sup_{\mathbf{v} \in W_{V_{k+j+1}}} \frac{b(\mathbf{v}, \mathbf{q})}{\|\mathbf{v}\|} \geq \mu \|\nabla \mathbf{q}\|, \quad \forall \mathbf{q} \in P_{k+j},$$

which implies the standard one [24].

For every fixed $\mathbf{q} \in P_{k+j}$ we want to construct $\mathbf{v} \in W_{V_{k+j+1}}$ such that

$$-\int_{\Omega} \mathbf{v} \cdot \nabla \mathbf{q}dx \geq c_1 \|\nabla \mathbf{q}\|^2, \quad (A.2)$$

$$\|\mathbf{v}\|_{0,\Omega} \leq c_2 \|\nabla \mathbf{q}\|. \quad (A.3)$$

Define:

$$\mathbf{v}(x, y) = (\hat{v}_1(x, y), \hat{v}_2(x, y)),$$

$$\hat{v}_1(x, y)_{\mu} = -\lambda_{AB}^a \lambda_{AE}^a \nabla \mathbf{q}\|L_{k-1,x}^a \cdot sign(H_a),$$
\begin{align*}
\hat{v}_2(x, y)_{la} &= -\lambda^a_{AB}\lambda^a_{AE} \frac{\partial \hat{q}}{\partial y}, \\
\hat{v}_1(x, y)_{lb} &= -\lambda^b_{ED}\lambda^b_{BD} \| \nabla \hat{q} \| L^{b}_{k+d-1,x} \cdot \text{sign}(H_b) - \lambda^b_{ED}\lambda^b_{EB} \frac{\partial \hat{q}}{\partial x}, \\
\hat{v}_2(x, y)_{rb} &= -\lambda^b_{ED}\lambda^b_{BD} \frac{\partial \hat{q}}{\partial y} - \lambda^b_{EB}\lambda^b_{ED} \| \nabla \hat{q} \| L^{b}_{k+d-1,y} \cdot \text{sign}(K_b), \\
\hat{v}_1(x, y)_{lc} &= -\lambda^c_{BC}\lambda^c_{CD} \frac{\partial \hat{q}}{\partial x}, \\
\hat{v}_2(x, y)_{rc} &= -\lambda^c_{BC}\lambda^c_{CD} \| \nabla \hat{q} \| L^{c}_{k+d-1,y} \cdot \text{sign}(K_c),
\end{align*}

where \text{sign}(x) is sign function defined as
\[
\text{sign}(x) = \begin{cases} 
1, & x > 0, \\
0, & x = 0, \\
-1, & x < 0.
\end{cases}
\]

and
\[
H_a = \int_{a} \lambda^a_{AB}\lambda^a_{AE} L_{k+d-1,x} \cdot \frac{\partial \hat{q}}{\partial y}, \quad H_b = \int_{b} \lambda^b_{ED}\lambda^b_{BD} L_{k+d-1,x} \cdot \frac{\partial \hat{q}}{\partial x}, \\
K_a = \int_{b} \lambda^b_{ED}\lambda^b_{ED} L_{k+d-1,y} \cdot \frac{\partial \hat{q}}{\partial y}, \quad K_b = \int_{c} \lambda^c_{BC}\lambda^c_{CD} L_{k+d-1,y} \cdot \frac{\partial \hat{q}}{\partial y}.
\]

First of all, we observe that \( \hat{v} \) is an element of \( W^{k+j+1} \), by the virtue of the fact that the tangential components of \( \nabla \hat{q} \) along the interface \( EB \) and \( BD \) are continuous.

It is easy to verify that \( \hat{v} \) satisfies (A.3). In order to check the validity of (A.2), define
\[
\| \nabla \hat{q} \|^2 = -\int_{\Omega} \hat{v} \cdot \nabla \hat{q} \cdot \cdot \cdot \cdot
\]

We verify that the expression \( \| \frac{\partial \hat{q}}{\partial x} \|_{H} : = |H_a| + |H_b| \) is a norm of \( \frac{\partial \hat{q}}{\partial x} \) in \( a \cup b \) and
\( \| \frac{\partial \hat{q}}{\partial y} \|_{K} : = |K_a| + |K_b| \) is a norm of \( \frac{\partial \hat{q}}{\partial y} \) in \( b \cup c \).

Step 1. We will show \( |H_a| + |H_b| \) vanishes only when \( \frac{\partial \hat{q}}{\partial x} \) equals zero. From (A.1)
\[
0 = \left\| \frac{\partial \hat{q}}{\partial x} \right\|_{H} = \left| \int_{x_{A}}^{0} L_{k+d-1,x} \cdot \frac{\partial \hat{q}}{\partial x} \right| + \int_{0}^{1} L_{k+d-1,x} \cdot k \frac{\partial \hat{q}}{\partial x}.
\]

From the orthogonality of Legendre polynomials \( L^{a}_{i,x}, L^{b}_{i,x}, \) and noting that \( \frac{\partial \hat{q}}{\partial x} \) is a homogeneous polynomial of degree \( k + d - 1 \), we have \( \frac{\partial \hat{q}}{\partial x} = 0 \) in \( a \cup b \).

Step 2. We will get \( \| k \frac{\partial \hat{q}}{\partial x} \|_{H} = |k| \| \frac{\partial \hat{q}}{\partial x} \|_{H} \) from
\[
\left\| k \frac{\partial \hat{q}}{\partial x} \right\|_{H} = \left| \int_{x_{A}}^{0} L_{k+d-1,x} \cdot k \frac{\partial \hat{q}}{\partial x} \right| + \int_{0}^{1} L_{k+d-1,x} \cdot k \frac{\partial \hat{q}}{\partial x}.
\]
Step 3. We will show that \( \| \hat{q}_1 \|_H + \| \hat{q}_2 \|_H \leq \| \hat{q}_1 \|_H + \| \hat{q}_2 \|_H \).

\[
\left\| \frac{\partial \hat{q}_1}{\partial x} + \frac{\partial \hat{q}_2}{\partial x} \right\|_H = \left| \int_{x_A}^0 \left( L_{k+d-1,x} \cdot \left( \frac{\partial \hat{q}_1}{\partial x} + \frac{\partial \hat{q}_2}{\partial x} \right) \right) \right| + \left| \int_0^1 L_{k+d-1,x} \cdot \left( \frac{\partial \hat{q}_1}{\partial x} + \frac{\partial \hat{q}_2}{\partial x} \right) \right|
\leq \left| \int_{x_A}^0 \left( L_{k+d-1,x} \cdot \frac{\partial \hat{q}_1}{\partial x} \right) \right| + \left| \int_0^1 L_{k+d-1,x} \cdot \frac{\partial \hat{q}_2}{\partial x} \right|
\leq \left| \int_0^1 L_{k+d-1,x} \cdot \frac{\partial \hat{q}_1}{\partial x} \right| + \left| \int_0^1 L_{k+d-1,x} \cdot \frac{\partial \hat{q}_2}{\partial x} \right|
= \left\| \frac{\partial \hat{q}_1}{\partial x} \right\|_H + \left\| \frac{\partial \hat{q}_2}{\partial x} \right\|_H.
\]

Similarly, \( \| \hat{q}_2 \|_K = |K_b| + |K_c| \) is a norm of \( \frac{\partial q}{\partial y} \) in \( b \cup c \). From the equivalence of norms on a finite dimensional space, there exists a constant \( c_A, c_B, c_C, c_H, c_K > 0 \) such that

\[
\int_a^\lambda \int_B^{\lambda q} \left( \frac{\partial \hat{q}}{\partial y} \right)^2 \geq c_A \int_a^\lambda \left( \frac{\partial \hat{q}}{\partial x} \right)^2, \quad \left\| \frac{\partial \hat{q}}{\partial y} \right\|_H \geq c_H \sqrt{\int_a^\lambda \left( \frac{\partial \hat{q}}{\partial x} \right)^2} + \int_b \left( \frac{\partial \hat{q}}{\partial x} \right)^2,
\]

\[
\int_b^\lambda \left( \frac{\partial \hat{q}}{\partial y} \right)^2 + \int_c^\lambda \left( \frac{\partial \hat{q}}{\partial y} \right)^2 \geq c_B \int_b \left( \frac{\partial \hat{q}}{\partial x} \right)^2 + \int_c \left( \frac{\partial \hat{q}}{\partial y} \right)^2,
\]

\[
\int_a^\lambda \int_B^{\lambda q} \left( \frac{\partial \hat{q}}{\partial y} \right)^2 \geq c_C \int_a^\lambda \left( \frac{\partial \hat{q}}{\partial x} \right)^2, \quad \left\| \frac{\partial \hat{q}}{\partial y} \right\|_H \geq c_K \sqrt{\int_b \left( \frac{\partial \hat{q}}{\partial y} \right)^2}.
\]

Set \( c_1 = \min\{c_A, c_B, c_C, c_H, c_K\} \) and obtain (A.2).

**Case** \( d = 3 \): We use the macroelement described by Stenberg in [23] in order to check the inf-sup condition. Let \( \mathcal{M} \) be a macroelement partition of the domain decomposition of \( T \). For a macroelement \( M \in \mathcal{M} \) we introduce the following usual notation:

\[
WV_M = \{ \hat{v} \in WV_{k+d+1} \cap \{ H^1 \} \}, \quad WP_M = \{ \hat{q} \in WP_{k+d} \}.
\]

Consider a generic macroelement \( M \in \mathcal{M} \). Let \( T_0 \in T \) be a tetrahedron of \( M \) and denote by \( x_0 \) the internal vertex of \( T_0 \) which also belongs to the other element of \( M \). There are three edges \( e_i, i = 1, \ldots, 3 \) of \( T_0 \) meeting at \( x_0 \). Thanks to the fact that \( x_0 \) is internal, none of the edges \( e_i \) lie on the boundary \( \partial \Omega \).

Let \( \hat{q} \in WP_M \) be given and suppose that

\[
\int_M \hat{q} \nabla \cdot \hat{v} = 0, \quad \forall \hat{v} \in WV_M.
\]  

We shall prove that \( \nabla \hat{q} \) vanishes on \( T_0 \), thus obtaining H1 condition described in Theorem 2.1 in [9] by virtue of the fact that \( T_0 \) is arbitrary and \( q \) is continuous.

First, we concentrate our attention on the edge \( e_1 \) and fix an \((x, y, z)\)-coordinate system in such a way that \( e_1 \) lies in the direction of the \( x \)-axis. We consider the collection \( \mathcal{A} = \{ T_0, \ldots, T_n \} \) of those elements of \( T \) which share the edge \( e_1 \) in common with \( T_0 \) (including \( T_0 \) itself). It is clear that \( T_i \in M \) and that exactly two faces of \( T_i \) touch other elements of \( \mathcal{A} \) of every \( i \).

Define \( \hat{v} \) in the following way:

\[
\hat{v}_{|T_i} = (\lambda_i \kappa_i \frac{\partial \hat{q}}{\partial x}, 0, 0),
\]
\( \hat{v}|_T = (0, 0, 0), \) if \( T \in \mathcal{T}, \; T \neq T_i, \; \forall i, \)

where \( \lambda_i \) and \( \kappa_i \) are the equations of the two faces of \( T_i \) which are not in common with any other element of \( \mathcal{A} \), normalized in order to assume the same value at the opposite vertex. It is worthwhile to observe that these two vertices are \( x_0 \) and the other extreme of the edge \( e_1 \).

It is easily verified that \( \hat{v} \) is a polynomial of degree \( k + 1 \) and that it is continuous in \( M \).

The continuity of \( \hat{q} \) in \( M \) ensures that the gradient of \( \hat{q} \) is continuous between two elements in all the directions which are contained in the plane of the interface; the \( x \)-axis is the direction of \( e_1 \) which is the edge of all common faces among the elements of \( \mathcal{A} \). Moreover, \( \hat{v} \) vanishes at the boundary of \( M \); hence, the following inclusion holds:

\[ \hat{v} \in W^{V_M}. \]

Suppose now that (A.5) hold.

\[ 0 = \int_M \hat{q} \nabla \hat{v} = -\int_M \nabla \hat{q} \cdot \hat{v} = -\sum_{i=1}^{n} \int_{T_i} \lambda_i \kappa_i \left( \frac{\partial \hat{q}}{\partial x} \right)^2. \]

It follows that the component of \( \nabla \hat{q} \) in the direction of the \( x \)-axis vanishes in \( T_i \) for every \( i \).

The same argument applies to the edge \( e_2 \) and \( e_3 \), giving the result that \( \nabla \hat{q} \) vanishes on \( T_0 \) in the direction of \( e_i \), for \( i = 1, \cdots, 3 \). These three directions being independent, the final result

\[ \nabla \hat{q} = (0, 0, 0), \] in \( T_0 \)

is obtained and the lemma is proved. Then the H1 condition of Theorem 2.1 in [9] is proved and the H2-H3 conditions follow immediately from the regularity assumption of \( T \).

\[ \square \]

References

1. Antonietti, P.F., da Veiga, L.B., Lovadina, C., Verani, M.: Hierarchical A Posteriori Error Estimators for the Mimetic Discretization of Elliptic Problems. SIAM J. Numer. Anal. 51(1), 654–675 (2013)
2. Araya, R., Barrenechea, G.R., Poza, A.: An adaptive stabilized finite element method for the generalized Stokes problem. J. Comput. Appl. Math. 214(2), 457–479 (2008)
3. Araya, R., Poza, A.H., Stephan, E.P.: A hierarchical a posteriori error estimate for an advection-diffusion-reaction problem. Math. Models Methods Appl. Sci. 15(7), 1119–1139 (2005)
4. Araya, R., Poza, A.H., Valentín, F.: On a hierarchical error estimator combined with a stabilized method for the Navier-Stokes equations. Numerical Methods for Partial Differential Equations 28(3), 782–806 (2012)
5. Araya, R., Rebolledo, R.: An a posteriori error estimator for a LPS method for Navier-Stokes equations. Appl. Numer. Math. 127, 179–195 (2018)
6. Arnold, D.N.: Spaces of Finite Element Differential Forms. In: Brezzi, F., Colli-Franzone, P., Gianazza, U.P., Gilardi, G. (eds.) Analysis and Numerics of Partial Differential Equations, Springer INdAM Series, vol. 4, Springer, Milan (2013)
7. Babuška, I., Aziz, A.: Survey lectures on the mathematical foundations of the finite element method, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations. Academic Press, New York, NY (1972)
8. Boffi, D.: Stability of higher order triangular Hood-Taylor methods for the stationary Stokes equations. Math. Models Methods Appl. Sci. 04(2), 223–235 (1994)
9. Boffi, D.: Three-Dimensional Finite Element Methods for the Stokes Problem. SIAM J. Numer. Anal. 34(2), 664–670 (1997)
10. Boffi, D., Brezzi, F., Demkowicz, L.F., Durán, R.G., Falk, R.S., Fortin, M.: Mixed Finite Elements, Compatibility Conditions, and Applications. Lecture Notes in Mathematics, vol. 1939. Springer, Berlin Heidelberg, Berlin, Heidelberg (2008)
11. Bornemann, F.A., Erdmann, B., Kornhuber, R.: A posteriori error estimates for elliptic problems in two and three space dimensions. SIAM J. Numer. Anal. 33(3), 1188–1204 (1996)
12. Dörfler, W.: A Convergent Adaptive Algorithm for Poisson’s Equation. SIAM J. Numer. Anal. 33(3), 1106–1124 (1996)
13. Dörfler, W., Nochetto, R.H.: Small data oscillation implies the saturation assumption. Numer. Math. 91(1), 1–12 (2002)
14. Eijkhout, V., Vassilevski, P.: The Role of the Strengthened Cauchy-Buniakowskii-Schwarz Inequality in Multilevel Methods. SIAM Rev. 33(3), 405–419 (1991)
15. Giani, S., Grubišić, L., Hakula, H., Ovall, J.S.: A Posteriori Error Estimates for Elliptic Eigenvalue Problems Using Auxiliary Subspace Techniques. Journal of Scientific Computing 88(3) (2021)
16. Hain, S., Ohlberger, M., Radic, M., Urban, K.: A hierarchical a posteriori error estimator for the Reduced Basis Method. Adv. Comput. Math. 45, 2191–2214 (2019)
17. Hakula, H., Neilan, M., Ovall, J.S.: A Posteriori Estimates Using Auxiliary Subspace Techniques. J. Sci. Comput. 72(1), 97–127 (2017)
18. Huang, W., Russell, R.D.: Adaptive Moving Mesh Methods, Applied Mathematical Sciences, vol. 174. Springer, New York (2011)
19. Logg, A., Mardal, K.A., Wells, G.: Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book (Lecture Notes in Computational Science and Engineering). Manual 84, 718 (2012)
20. Nochetto, R.H., Siebert, K.G., Veeser, A.: Theory of adaptive finite element methods: An introduction. Multiscale, Nonlinear and Adaptive Approximation: Dedicated to Wolfgang Dahmen on the Occasion of his 60th Birthday pp. 409–542 (2009)
21. Pfeiler, C.M., Praetorius, D.: Dörfler marking with minimal cardinality is a linear complexity problem. Math. Comput. 89(326), 2735–2752 (2020)
22. Scott, L.R., Zhang, S.: Finite Element Interpolation of Nonsmooth Functions Satisfying Boundary Conditions. Math. Comput. 54(190), 483 (1990)
23. Stenberg, R.: On some three-dimensional finite elements for incompressible media. Comput. Methods Appl. Mech. Eng. 63(3), 261–269 (1987)
24. Verfürth, R.: Error estimates for a mixed finite element approximation of the Stokes equations. RAIRO. Analyse numérique 18(2), 175–182 (1984)
25. Verfürth, R.: A review of a posteriori error estimation and adaptive mesh-refinement techniques. Teubner, Stuttgart (1996)
26. Verfürth, R.: A Posteriori Error Estimation Techniques for Finite Element Methods. Oxford University Press, Oxford (2013)
27. Zienkiewicz, O., Craig, A.: Adaptive refinement, errorestimates, multigrid solution and hierarchical finite element method concepts, pp. 25–59. Accuracy Estimates and Adaptive Refinements in Finite Element Computations, John Wiley and Sons pp (1986)
28. Zienkiewicz, O., Kelley, D., Gago, J., Babuška, I.: Hierarchical finite element approaches, error estimates and adaptive refinement, pp. 313–346. The Mathematics of Finite Elements and Applications IV, Academic Press pp (1982)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.