Convergence of an adaptive mixed finite element method for convection-diffusion-reaction equations

DU ShaoHong\textsuperscript{1,2} & XIE XiaoPing\textsuperscript{3,}\textsuperscript{*}

\textsuperscript{1}School of Science, Chongqing Jiaotong University, Chongqing 400074, China;
\textsuperscript{2}Beijing Computational Science Research Center, Beijing 100084, China;
\textsuperscript{3}School of Mathematics, Sichuan University, Chengdu 610064, China

Email: duzheyan.student@sina.com, xpxie@scu.edu.cn

Received November 17, 2013; accepted January 22, 2015; published online March 17, 2015

Abstract  We prove the convergence of an adaptive mixed finite element method (AMFEM) for (nonsymmetric) convection-diffusion-reaction equations. The convergence result holds for the cases where convection or reaction is not present in convection- or reaction-dominated problems. A novel technique of analysis is developed by using the superconvergence of the scalar displacement variable instead of the quasi-orthogonality for the stress and displacement variables, and without marking the oscillation dependent on discrete solutions and data. We show that AMFEM is a contraction of the error of the stress and displacement variables plus some quantity. Numerical experiments confirm the theoretical results.

Keywords  convection-diffusion-reaction equation, adaptive mixed finite element method, superconvergence, oscillation, convergence

MSC(2010) 65N30, 65N15, 65N12, 65N50

Citation: Du S H, Xie X P. Convergence of an adaptive mixed finite element method for convection-diffusion-reaction equations. Sci China Math, 2015, 58: 1327–1348, doi: 10.1007/s11425-015-4992-6

1 Introduction and main results

Let $\Omega$ be a bounded polygonal or polyhedral domain in $\mathbb{R}^d$, $d = 2$ or $3$. We consider the following convection-diffusion-reaction equations:

$$
\begin{aligned}
-\nabla \cdot (S \nabla p) + \nabla \cdot (pw) + rp &= f, \\
p &= 0,
\end{aligned}
\quad \text{in } \Omega,
\quad \text{on } \partial \Omega,
$$

(1.1)

where $S \in L^\infty(\Omega; \mathbb{R}^{d\times d})$ is an inhomogeneous and anisotropic diffusion-dispersion tensor, $w$ is a (dominating) velocity field, $r$ is a reaction function, and $f$ is a source term. The choice of homogeneous boundary condition is made for ease of presentation, since similar results are valid for other boundary conditions.

Adaptive methods for the numerical solutions of PDEs are now standard tools in science and engineering to achieve better accuracy with minimum degrees of freedom. The adaptive procedure of (1.1) consists of loops of the form

SOLVE $\rightarrow$ ESTIMATE $\rightarrow$ MARK $\rightarrow$ REFINE.

(1.2)
A posteriori error estimation (ESTIMATE) is an essential ingredient of adaptivity, and reaches its mature level after two decades of development [1, 4–6, 12, 13, 22, 29, 41, 43, 44]. However, the analysis of convergence of the whole algorithm (1.2) is still in its infancy, and is carried out mainly for standard adaptive finite element methods (AFEM) [15, 19, 33–35].

Due to the saddle-point characteristic of the mixed finite element approximation, there is no orthogonality available, as is one of the main difficulties in the convergence analysis of AMFEM. Thus one has to find some quasi-orthogonality instead of the orthogonality, and the occurrence of oscillation of data is inevitable. Hence, how to deal with the oscillation becomes a key issue in the analysis. For the convergence of AMFEM, the present studies mainly focus on Poisson equations. In [11], Carstensen and Hoppe proved the error reduction and convergence for the lowest-order Raviart-Thomas element with marking the data oscillation. Chen et al. [16] showed the convergence of a quasi-error with marking the data oscillation. In [7, 14, 23], the convergence was analyzed for the lowest-order Raviart-Thomas element where the local refinement was performed by using only either the estimators or the data oscillation term.

For general diffusion problems and more general mixed elements, by using the orthogonality of the divergence of the flux, Du and Xie [24] showed the convergence of the flux error plus some quantity without marking the oscillation.

The purpose of this paper is to prove the following convergence results for an AMFEM for the convection-diffusion-reaction equations (1.1) and verify them computationally.

**Theorem 1.1** (Convergence of AMFEM). Denote by \( \{T_k, e_k, A_k^2\}_{k \geq 0} \) the sequence of meshes, the error of the stress and displacement variables, and some quantity (defined by (5.12)) produced by the AMFEM algorithm. Let \( h_0 \) be the mesh size of the quasi-uniform initial mesh \( T_0 \), with the assumptions (D1)–(D6) of data and the weak regularity (3.3) on \( S \). Then there exist two positive constants \( q \) and \( \alpha \in (0,1) \) such that

\[
e_k^2 + (1 - h_0 q)A_k^2 \leq \alpha^2 (e_0^2 + (1 - h_0 q)A_0^2)
\]

when \( h_0 \leq \frac{1 - \alpha^2}{1 + \alpha^2} q \). This means that AMFEM, as \( h_0 \) is small enough, converges with a linear rate \( \alpha \), namely,

\[
e_k^2 + (1 - h_0 q)A_k^2 \leq \alpha^k (e_0^2 + (1 - h_0 q)A_0^2).
\]

This theorem extends the convergence results in [24] in the following several aspects.

- We deal with more general convection-diffusion-reaction equations here with variable coefficients \( S, w \) and \( r \), whereas in [24] \( w \) and \( r \) vanish.
- The orthogonality for the divergence of the flux is absent due to the convection term \( w \cdot \nabla p \) and the zero order term \((r + \nabla \cdot w)p\). So this contribution considers not only the flux (stress variable) error but also the displacement variable error.
- The quasi-orthogonality for stress and displacement variables also fails due to the terms \( w \cdot \nabla p \) and \((r + \nabla \cdot w)p\). We develop a new technique of analysis based on the superconvergence of the scalar displacement variable instead of the quasi-orthogonality for the stress and displacement variables. This will lead to an additional constraint on the mesh size, \( h_0 \), of the quasi-uniform initial mesh \( T_0 \).
- The oscillation term depends on the discrete solution and data. Therefore, the oscillation and error cannot be reduced separately here. In [24], the oscillation term is not included in the a posteriori indicators.
- Since the error and oscillation are now coupled, in order to prove convergence without marking the oscillation term, we need to handle them together by following the same idea as in [18, 28].
- In comparison with the previous analysis methods, we consider the a posteriori indicators with weighted factors. We also release the constraint that the divergence of the convection term is free in contrast to the analysis of the standard AFEM (see [30]).

The rest of this paper is organized as follows. Section 2 gives some preliminaries and details on notations. Section 3 derives an estimate for the error between the \( L^2 \)-projection of the displacement and its approximation solution, which is key to the convergence analysis. Section 4 shows the estimator reduction. We prove Theorem 1.1 (Convergence of AMFEM algorithm) in Section 5 and present four numerical experiments to illustrate properties of the AFMEM in Section 6.
2 Assumptions, weak problem, and AMFEM algorithm

For a domain $A \subset \mathbb{R}^d$, we denote by $L^2(A)$ and $L^2(A) := (L^2(A))^d$ the spaces of square-integrable functions, by $(\cdot , \cdot)_A$ the $L^2(A)$ or $L^2(A)$ inner product, by $\| \cdot \|_A$ the associated norm, and by $|A|$ the Lebesgue measure of $A$. Let $H^k(A)$ be the usual Sobolev space equipped with the norm $\| \cdot \|_{k,A}$ for $k = 1, 2$:

$$H^1(A) := \{ v \in H^{1}(A) : v|_{\partial A} = 0 \} \quad \text{and} \quad H(\text{div}, A) := \{ v \in L^2(A) : \text{div } v \in L^2(A) \}.$$ 

$\langle \cdot , \cdot \rangle_{\partial A}$ denotes the $(d-1)$-dimensional inner product on $\partial A$ for the duality pairing between $H^{-1/2}(\partial A)$ and $H^{1/2}(\partial A)$. In what follows we shall omit the subscript $\Omega$ when $A = \Omega$.

Let $T_h$ be a shape regular triangulation in the sense of [17], and denote the mesh size by $h_T := |T|^{1/d}$ with $|T|$ being the volume of $T \in T_h$. Let $C_Q$ be a positive constant depending only on a quantity $Q$, and $C_i(i = 1, 2, \ldots)$ positive constants determined only by the shape regularity of $T_h$. We denote by $\varepsilon_h$ the set of element sides in $T_h$, by $\varepsilon_h^0$ the set of interior sides of elements. For $K \in T_h$, denote by $\varepsilon_K$ the set of sides of $K$. Furthermore, we denote by $\omega_K$ and $\omega_\sigma$ the unions of all elements in $T_h$ respectively sharing a side with $K$ and sharing a side $\sigma \in \varepsilon_h$. We use the “broken Sobolev space”

$$H^1(\bigcup T_h) := \{ \varphi \in L^2(\Omega) : \varphi|_K \in H^1(K), \forall K \in T_h \}.$$ 

$H^2(\bigcup \bigcup T_h)$ is defined analogously. Denote by $[v]|_\sigma := ([v]|_\sigma - (v|_K)|_\sigma$ the jump of $v$ in $H^1(\bigcup T_h)$ over an interior side $\sigma := K \cap L$ of diameter $h_\sigma := \text{diam}(\sigma)$, shared by the two neighboring (closed) elements $K, L \in T_h$. Especially, $[v]|_\sigma := ([v]|_\sigma$ if $\sigma \subset \partial K \cap \partial \Omega$. Note that $[\cdot]$ is a linear operator over the broken Sobolev space $H^1(\bigcup T_h)$.

We note that throughout the paper, the local version of the differential operator $\nabla$ is denoted by $\nabla_h$, namely, $\nabla_h : H^1(\bigcup T_h) \to (L^2(\Omega))^d$ is defined by $\nabla_h v|_K := \nabla(v|_K)$ for all $K \in T_h$.

Given a unit normal vector $n_\sigma = (n_1, \ldots, n_d)^T$ along the side $\sigma$ with $d = 2, 3$, we define the tangential component of a vector $v \in \mathbb{R}^d$ with respect to $n_\sigma$ by

$$\gamma_{t_\sigma}(v) := \begin{cases} v \cdot (-n_2, n_1)^T, & \text{if } d = 2; \\ v \times n_\sigma, & \text{if } d = 3, \end{cases}$$

where $\times$ denotes the usual vector product of two vectors in $\mathbb{R}^3$.

Following [42], we suppose that there exists an original triangulation $T_0$ of $\Omega$ such that data of the problem (1.1) is given in the following way.

**Assumptions of data:**

(D1) $S_K := S|_K$ is a constant, symmetric, and uniformly positive definite tensor such that $c_{S,K} v \cdot v \leq S_K v \cdot v \leq C_{S,K} v \cdot v$ holds for all $v \in \mathbb{R}^d$ and all $K \in T_0$ with $c_{S,K} > 0, C_{S,K} > 0$;

(D2) $w \in RT_0(T_0)$ (see below) and $|w|_K \leq C_{w,K}$ for all $K \in T_0$ with $C_{w,K} \geq 0$;

(D3) $r_K := r|_K$ is a constant for all $K \in T_0$;

(D4) $c_{w,r,K} := 1/2 \nabla \cdot w|_K + r|_K \geq 0$ and $C_{w,r,K} := |\nabla \cdot w|_K + r|_K|$ for all $K \in T_0$;

(D5) $f \in L^2(\Omega)$;

(D6) if $c_{w,r,K} = 0$, then $C_{w,r,K} = 0$.

Note that in [21, 25] $f|_K$ is assumed to be a polynomial of degree at most $k$ for each $K \in T_0$ so as to derive the efficiency of the residual indicators. Here we relax the restriction of $f$ (see (D5)).

Introduce the stress variable $u := -S \nabla p$, the mixed variational problem of (1.1) reads as: Find $(u, p) \in H(\text{div}, \Omega) \times L^2(\Omega)$ such that

$$\begin{align*}
(S^{-1} u, v) - (p, \nabla \cdot v) &= 0, \quad \text{for all } v \in H(\text{div}, \Omega), \\
(\nabla \cdot u, \varphi) - (S^{-1} u \cdot w, \varphi) + ((r + \nabla \cdot w)p, \varphi) &= (f, \varphi), \quad \text{for all } \varphi \in L^2(\Omega).
\end{align*}$$

(2.1)

(2.2)

Let $P_0(K)$ denote the set of constant functions on each $K \in T_h$. We respectively define the lowest order Raviart-Thomas finite element (see [36]) space and the piecewise constant space as following:

$$RT_0(T_h) := \left\{ q_h \in H(\text{div}, \Omega) : \forall K \in T_h, \exists a \in \mathbb{R}^d, \exists b \in \mathbb{R}, \text{such that } q_h(x) = a + bx, \text{for all } x \in K \right\},$$

$$P_0(T_h) := \{ q_h \in L^2(\Omega) : \forall K \in T_h, q_h|_K \text{ is constant on each } K \in T_h \}.$$
We note that \( \nabla \cdot (RT_0(T_h)) \subset P_0(T_h) \).

The centered mixed finite element scheme (see [20, 42]) of (1.1) reads as: Find \((u_h, p_h) \in RT_0(T_h) \times P_0(T_h)\) such that

\[
\begin{align*}
(S^{-1}u_h, v_h) - (p_h, \nabla \cdot v_h) &= 0, \quad \text{for all } v_h \in RT_0(T_h), \\
(\nabla \cdot u_h, \varphi_h) - (S^{-1}u_h \cdot w, \varphi_h) + ((r + \nabla \cdot w)p_h, \varphi_h) &= (f, \varphi_h), \quad \text{for all } \varphi_h \in P_0(T_h).
\end{align*}
\]

(2.3)

(2.4)

In what follows, we shall show an AMFEM algorithm based on the a posteriori error estimator developed in [25]. We note that our convergence analysis below is also valid for the AMFEM based on the estimator proposed in [21].

Suppose that the module SOLVE outputs a pair of discrete solutions over \( T_h \), namely, \((u_h, p_h) = \text{SOLVE}(T_h)\). The estimator in [25] consists of several indicators with different weight factors, where the elementwise estimator \( \eta^2_{T_h}(u_h, p_h, K) \) can, for convenience, be abbreviated to

\[
\eta^2_{T_h}(u_h, p_h, K) := D^2_K h^2_K \|S^{-1}u_h\|^2_K + \alpha^2_K h^2_K \|R_K\|^2_K + \sum_{\sigma \in \varepsilon_K} D^2_\sigma h_\sigma \|\gamma_{\varepsilon_{T_h}}(S^{-1}u_h)\|^2_\sigma.
\]

Here

\[
\alpha_K = \min(h_K/\sqrt{c_{S,K}}, 1/\sqrt{c_{w,r,K}}),
\]

\(R_K\) is the elementwise residual defined by

\[
R_K := f - \nabla \cdot u_h + (S^{-1}u_h) \cdot w - (r + \nabla \cdot w)p_h,
\]

and \(D_K, D_\sigma\) denote two variants of coefficients over each element \( K \in T_h \) and each side \( \sigma \in \varepsilon_K \) respectively given by

\[
D^2_K := c_{w,r,K} + C_{w,r,K}^2 \alpha^2_K + \max_{K' \cap K \neq \emptyset} \frac{c_{w,r,K'}}{c_{w,r,K}} + \max_{K' \cap K \neq \emptyset} \frac{\|\nabla \cdot w\|^2_{K'}}{C_{w,r,K}^2},
\]

\[
D^2_\sigma := \frac{1}{2} \max_{K : K \cap \sigma \neq \emptyset} C_{S,K} + \frac{1}{2} \min \left\{ \max_{K : K \cap \sigma \neq \emptyset} \frac{C_{w,K}}{c_{w,r,K}}, \max_{K : K \cap \sigma \neq \emptyset} \frac{h^2_K C_{w,K}}{c_{S,K}} \right\}.
\]

Define the global and local errors, \( e_h \) and \( E_K \), of the stress and displacement variables as

\[
e^2_h := \sum_{K \in T_h} \varepsilon^2_K, \quad \varepsilon^2_K := \|S^{-1/2}(u - u_h)\|^2_K + c_{w,r,K} \|p - p_h\|^2_K.
\]

(2.5)

From [25], or [21] but with different forms of \(D_K\) and \(D_\sigma\), an upper bound estimate holds as follows,

\[
e^2_h \leq C_1 \eta^2_{T_h}(u_h, p_h, T_h) := C_1 \sum_{K \in T_h} \eta^2_{T_h}(u_h, p_h, K),
\]

(2.6)

where the positive constant \(C_1\) depends only on the shape regularity of the meshes.

For a given triangulation \( T_h \) and a pair of corresponding discrete solutions \((u_h, p_h) \in RT_0(T_h) \times P_0(T_h)\), we assume that the module ESTIMATE outputs the indicators

\[
\{\eta^2_{T_h}(u_h, p_h, K)\}_{K \in T_h} = \text{ESTIMATE}(u_h, p_h, T_h).
\]

Let \( \bar{R}_K \) denote the mean of \( R_K \) over each element \( K \in T_h \). We define the oscillation

\[
\text{osc}^2_h := \sum_{K \in T_h} h^2_K \|R_K - \bar{R}_K\|^2_K.
\]

(2.7)

We note that throughout this paper the triangulation \( T_h \) means a refinement of \( T_H \), and all notation with respect to the mesh \( T_H \) are defined similarly. We shall also use the notation \( A \lesssim B \) to represent \( A \leq CB \) with \( C > 0 \) being a mesh-size independent, generic constant.
In the MARK step, by Dörfler marking we select the elements to mark according to the indicators, namely, given a grid $\mathcal{T}_H$ with the set of indicators $\{\eta^2_{\mathcal{T}_H}(u_H, p_H, K)\}_{K \in \mathcal{T}_H}$ and marking parameter $\theta \in (0, 1]$, the module MARK outputs a subset of marked elements, $\mathcal{M}_H \subset \mathcal{T}_H$, with
\[
\mathcal{M}_H = \text{MARK}(\{\eta^2_{\mathcal{T}_H}(u_H, p_H, K)\}_{K \in \mathcal{T}_H}, \mathcal{T}_H, \theta)
\]
satisfying Dörfler property
\[
\eta_{\mathcal{T}_H}(u_H, p_H, \mathcal{M}_H) := \left( \sum_{K \in \mathcal{M}_H} \eta^2_{\mathcal{T}_H}(u_H, p_H, K) \right)^{1/2} \geq \theta \eta_{\mathcal{T}_H}(u_H, p_H, \mathcal{T}_H).
\]

In the REFINE step, we suppose that the refinement rule, such as the longest edge bisection [37,38] or the newest vertex bisection [31,32,40], is guaranteed to produce conforming and shape regular meshes. Given a fixed integer $b \geq 1$, a mesh $\mathcal{T}_H$, and a subset $\mathcal{M}_H \subset \mathcal{T}_H$ of marked elements, a conforming triangulation $\mathcal{T}_H$ is output by
\[
\mathcal{T}_H = \text{REFINE}(\mathcal{T}_H, \mathcal{M}_H),
\]
where all the elements of $\mathcal{M}_H$ are at least bisected $b$ times. Note that not only the marked elements get refined but also the additional elements are refined to recover the conformity of triangulations.

We now describe the AMFEM algorithm. In doing so, we replace the subscript $H$ by an iteration counter called $k \geq 0$. Let $\mathcal{T}_0$ be a uniform triangulation with a marking parameter $\theta \in (0, 1]$. The basic loop of AMFEM is then given by the following iterations:

**Algorithm 2.1 (The AMFEM algorithm).** Set $k = 0$. Repeat (1)–(4) until a given number of iterations is reached.

1. $(u_k, p_k) = \text{SOLVE}(\mathcal{T}_k)$;
2. $\{\eta^2_k(u_k, p_k, K)\}_{K \in \mathcal{T}_k} = \text{ESTIMATE}(u_k, p_k, \mathcal{T}_k)$;
3. $\mathcal{M}_k = \text{MARK}(\{\eta^2_k(u_k, p_k, K)\}_{K \in \mathcal{T}_k}, \mathcal{T}_k, \theta)$;
4. $\mathcal{T}_{k+1} = \text{REFINE}(\mathcal{T}_k, \mathcal{M}_k)$; $k = k + 1$.

We note that the AMFEM algorithm is a standard one in which it employs only the error estimator $\{\eta^2_k(u_k, p_k, K)\}_{K \in \mathcal{T}_k}$ and needs neither marking the oscillation nor the interior node property.

## 3 Estimate for the $L^2$-projection of the displacement

This section is devoted to the estimation of $\|Q_h p - p_h\|$, where $Q_h$ is the $L^2$-projection operator onto $P_0(\mathcal{T}_h)$. The estimate is one key to the proof of convergence without the quasi-orthogonality available due to the convection term. It gives as well a posteriori error estimates for the $L^2$-projection of the displacement variable (see remark 3.1).

Consider the following auxiliary problem:
\[
\begin{cases}
\nabla \cdot (S \nabla \phi) + \nabla \phi \cdot w - r \phi = Q_h p - p_h, & \text{in } \Omega,

\phi = 0, & \text{on } \partial \Omega.
\end{cases}
\]
(3.1)

It is well known that there exists a unique solution $\phi \in H^1_0(\Omega)$ to the problem (3.1) when the convection and reaction terms satisfy $r + 1/2 \nabla \cdot w \geq 0$ (Assumptions (D1) and (D4)) with the following regularity estimate:
\[
\|\phi\|_{H^1} = \|\phi\|_{H^1(\bigcup \mathcal{T}_k)} \lesssim \|Q_h p - p_h\|.
\]
(3.2)

Moreover, if $\Omega$ is convex, $S \in C^{1,0}(\Omega)$ implies the estimate
\[
\|\phi\|_{H^2(\bigcup \mathcal{T}_k)} \lesssim \|Q_h p - p_h\|.
\]
(3.3)

We emphasize that we only need an estimate on $\|\phi\|_{H^2(K)}$ for each $K \in \mathcal{T}_h$, i.e., the assumption on $S$ could be weakened in the sense that only (3.3) is required. In [10], Carstensen gave an example which showed that when $S$ is a piecewise constant, $\phi$ satisfies (3.3) but is not $H^2$-regular.
Set
\[ z := S \nabla \phi \in H(\text{div}, \Omega) \cap (H^1(\cup T_h))^d. \]
Denote by \( \phi_h \) the \( L^2 \)-projection of \( \phi \) onto \( P_0(T_h) \), and \( \Pi_h \) the interpolation operator from \( H(\text{div}, \Omega) \cap (L^s(\Omega))^d \) (for fixed \( s > 2 \)) (see [9, Chapter III.3]) onto \( RT_0(T_h) \) with the following estimate:
\[ \| h^{-1}(z - \Pi_h z) \| \lesssim |z|_{H^1(\cup T_h)} \quad \text{for all} \quad z \in H(\text{div}, \Omega) \cap (H^1(\cup T_h))^d. \]
(3.4)
Recall the commuting property of the interpolation operator \( \Pi_h \) as follows:
\[ \nabla \cdot (\Pi_h z) = Q_h(\nabla \cdot z). \]
(3.5)
We refer to [3, 9, 27] for the detailed construction of such an interpolation operator \( \Pi_h \) and the approximation property.

From (2.1) and (2.3), we obtain
\[ (Q_h p - p_h, \nabla \cdot \Pi_h z) = (p - p_h, \nabla \cdot \Pi_h z) = (S^{-1}(u - u_h), \Pi_h z). \]
(3.6)
An integration by parts implies
\[ (S^{-1}(u - u_h), z) = (S^{-1}(u - u_h), S \nabla \phi) = - (\nabla \cdot (u - u_h), \phi). \]
(3.7)
From (2.2) and (2.4) it follows that
\[ (\nabla \cdot (u - u_h), \phi_h) = (S^{-1}(u - u_h) \cdot w, \phi_h) - ((r + \nabla \cdot w)(p - p_h), \phi_h) \\
= (S^{-1}(u - u_h) \cdot w, \phi_h) - (p - p_h, (r + \nabla \cdot w)\phi_h) \\
= (S^{-1}(u - u_h) \cdot w, \phi_h) - (Q_h p - p_h, (r + \nabla \cdot w)\phi_h). \]
(3.8)
Denote
\[ I := (Q_h(\nabla \phi \cdot w), Q_h p - p_h) - (r \phi_h, Q_h p - p_h). \]
In view of the commuting property the interpolation operator \( \Pi_h \), a combination of (3.6)–(3.8) yields
\[ \|Q_h p - p_h\|^2 = (Q_h p - p_h, Q_h \nabla \cdot z) + I = (Q_h p - p_h, \nabla \cdot \Pi_h z) + I \\
= (S^{-1}(u - u_h), \Pi_h z - z) + (S^{-1}(u - u_h), z) + I \\
= (S^{-1}(u - u_h), \Pi_h z - z) - (\nabla \cdot (u - u_h), \phi - \phi_h) - (\nabla \cdot (u - u_h), \phi_h) + I \\
= (S^{-1}(u - u_h), \Pi_h z - z) - (\nabla \cdot (u - u_h), \phi - \phi_h) - (S^{-1}(u - u_h) \cdot w, \phi_h) \\
+ (\nabla \cdot w \phi_h, Q_h p - p_h) + (Q_h(\nabla \phi \cdot w), Q_h p - p_h) \\
= (S^{-1}(u - u_h), \Pi_h z - z) - (\nabla \cdot (u - u_h), \phi - \phi_h) - (S^{-1}(u - u_h) \cdot w, \phi_h - \phi) \\
- (S^{-1}(u - u_h) \cdot w, \phi) + (\nabla \cdot w \phi_h, Q_h p - p_h) + (Q_h(\nabla \phi \cdot w), Q_h p - p_h). \]
(3.9)
Recall the postprocessed technique developed by Vohralík in [42], where a postprocessed approximation \( \tilde{p}_h \) to the displacement \( p \) is constructed such that
\[ -S_K \nabla \tilde{p}_h |_K = u_h \quad \text{and} \quad \frac{1}{|K|} \int_K \tilde{p}_h \, dx = p_h |_K, \quad \text{for all} \ K \in T_h. \]
Then, from \( w \in RT_0(T_h) \), we have
\[ -(S^{-1}(u - u_h) \cdot w, \phi) = \sum_{K \in T_h} \int_K \nabla (p - \tilde{p}_h) \cdot w \phi \\
= \sum_{K \in T_h} \int_K \nabla \cdot ((p - \tilde{p}_h) w) \phi - \nabla \cdot w (p - \tilde{p}_h) \phi \\
= \sum_{K \in T_h} \int_K -\nabla \phi \cdot w (p - \tilde{p}_h) + \int_{\partial K} (p - \tilde{p}_h) w \cdot n \phi - (\nabla \cdot w \phi, p - \tilde{p}_h) \]
Proof. For convenience, we use the following notations, and 

\[ (Q_h(\nabla \phi \cdot w), Q_h p - p_h) = (Q_h(\nabla \phi \cdot w), p - p_h) = (Q_h(\nabla \phi \cdot w), p - p_h) \]  

(3.11) 

and 

\[ (\nabla \cdot w \phi_h, Q_h p - p_h) = (\nabla \cdot w \phi_h, p - p_h) = (\nabla \cdot w \phi_h, p - p_h). \]  

(3.12) 

For convenience, we use the following notations, 

\[ I_1 := (S^{-1}(u - u_h), \Pi_h z - z), \quad I_2 := -(\nabla \cdot (u - u_h), \phi - \phi_h), \]  

\[ I_3 := -(S^{-1}(u - u_h) \cdot w, \phi - \phi_h), \quad I_4 := -(\nabla \phi \cdot w - Q_h(\nabla \phi \cdot w), p - p_h), \]  

\[ I_5 := -(\nabla \cdot w(\phi - \phi_h), p - p_h), \quad I_6 := -\sum_{\sigma \in e_h} \int_{\sigma} [\tilde{p}_h] w \cdot n. \]  

From (3.9)–(3.12) we arrive at 

\[ ||Q_h p - p_h||^2 = \sum_{i=1}^{4} I_i - (\nabla \cdot w \phi, p - p_h) + (\nabla \cdot w \phi_h, Q_h p - p_h) + I_6 \]  

(3.13) 

From (3.4) and (3.3) it follows that 

\[ I_1 = \sum_{K \in T_h} (S^{-1}(u - u_h), \Pi_h z - z)_K \]  

\[ \lesssim \sum_{K \in T_h} ||S^{-1/2}(u - u_h)||_K ||\Pi_h z - z||_K \]  

\[ \lesssim ||h||_{L^\infty(\Omega)} e_h ||z||_{H^1(\bigcup T_h)} \]  

\[ \lesssim ||h||_{L^\infty(\Omega)} e_h ||Q_h p - p_h||. \]  

(3.15) 

Lemma 3.1. Denote by \( ||h||_{L^\infty(\Omega)} \) the maximum norm of the mesh-size function \( h \) with respect to \( T_h \), \( e_h \) the error defined in (2.5). Then it holds that 

\[ I_1 \lesssim ||h||_{L^\infty(\Omega)} e_h ||Q_h p - p_h||. \]  

(3.14) 

Proof. From (3.4) and (3.3) it follows that 

\[ I_2 \lesssim (||h||_{L^\infty(\Omega)} e_h + \mathrm{osc}_h) ||Q_h p - p_h||. \]  

(3.16) 

Lemma 3.2. It holds that 

\[ I_1 \lesssim (||h||_{L^\infty(\Omega)} e_h + \mathrm{osc}_h) ||Q_h p - p_h||. \]  

(3.17) 

Proof. Notice that (2.2) can be equivalently written as: 

\[ (\nabla \cdot u, \varphi)_K - (S^{-1} u \cdot w, \varphi)_K + ((r + \nabla \cdot w)p, \varphi)_K = (f, \varphi)_K, \quad \text{for all } \varphi \in L^2(K), \quad K \in T_h. \]  

(3.17) 

Meanwhile, the relation (2.4) can be equivalently written as: 

\[ (\nabla \cdot u_h, \varphi)_K - (S^{-1} u_h \cdot w, \varphi)_K + ((r + \nabla \cdot w)p_h, \varphi)_K = (f, \varphi)_K, \quad \text{for all } \varphi \in P_0(K), \quad K \in T_h. \]  

(3.18)
For an arbitrary \( \varphi \in L^2(K) \), let \( \bar{\varphi}_K \) denote the mean of \( \varphi \) over \( K \in T_h \). A combination of (3.17) and (3.18) yields that

\[
(\nabla \cdot (u - u_h), \varphi)_K = (\nabla \cdot u, \varphi)_K - (\nabla \cdot u_h, \varphi - \bar{\varphi}_K)_K - (\nabla \cdot u_h, \bar{\varphi}_K)_K \\
= (R_K, \varphi - \bar{\varphi}_K)_K + (S^{-1}(u - u_h) \cdot \mathbf{w}, \varphi)_K - ((r + \nabla \cdot \mathbf{w})(p - p_h), \varphi)_K \\
= (R_K - \bar{R}_K, \varphi - \bar{\varphi}_K)_K + (S^{-1}(u - u_h) \cdot \mathbf{w}, \varphi)_K - ((r + \nabla \cdot \mathbf{w})(p - p_h), \varphi)_K \\
\leq \| R_K - \bar{R}_K \|_K + \| S^{-1}(u - u_h) \|_K \| \mathbf{w} \|_{L^\infty(K)} + C_{w,r,K} \| p - p_h \|_K \| \varphi \|_K.
\]

Here we recall that \( \bar{R}_K \) is the mean of the elementwise residual \( R_K \) over each \( K \in T_h \). The above inequality indicates

\[
\| \nabla \cdot (u - u_h) \|_K = \sup_{\varphi \in L^2(K), \varphi \neq 0} \frac{(\nabla \cdot (u - u_h), \varphi)_K}{\| \varphi \|_{L^2(K)}} \lesssim \| R_K - \bar{R}_K \|_K + \mathcal{E}_K. \tag{3.19}
\]

Then it follows that

\[
I_2 = -\sum_{K \in T_h} (\nabla \cdot (u - u_h), \phi - \phi_h) \\
\lesssim \sum_{K \in T_h} \| \nabla \cdot (u - u_h) \|_K h_K \| \nabla \phi \|_K \\
\lesssim \| h \|_{L^\infty(\Omega)} e_h \| \nabla \phi \|. \tag{3.20}
\]

The desired result (3.16) follows from (3.20) and (3.2).

**Lemma 3.3.** It holds that

\[
I_3 \lesssim \| h \|_{L^\infty(\Omega)} e_h \| Q_h p - p_h \|. \tag{3.21}
\]

**Proof.** By noticing

\[
I_3 = -\sum_{K \in T_h} (S^{-1}(u - u_h) \cdot \mathbf{w}, \phi_h - \phi)_K \\
\lesssim \sum_{K \in T_h} \| S^{-1}(u - u_h) \|_K h_K \| \nabla \phi \|_K \\
\lesssim \| h \|_{L^\infty(\Omega)} e_h \| \phi \|_{H^1}, \tag{3.22}
\]

the desired result (3.21) follows from (3.22) and the regularity estimate (3.2).

**Lemma 3.4.** It holds that

\[
I_4 \lesssim \| h \|_{L^\infty(\Omega)} e_h \| Q_h p - p_h \|. \tag{3.23}
\]

**Proof.** Recall a local efficiency estimate of \( h_K \| S^{-1}u_h \|_K \) as following (see Lemma 7.3 in [21,25]):

\[
h_K \| S^{-1}u_h \|_K \lesssim \mathcal{E}_K. \tag{3.24}
\]

By the triangle inequality we obtain

\[
I_4 = -\sum_{K \in T_h} (\nabla \phi \cdot \mathbf{w} - Q_h (\nabla \phi \cdot \mathbf{w}), p - \tilde{p}_h)_K \\
\lesssim \sum_{K \in T_h} h_K \| \nabla (\nabla \phi \cdot \mathbf{w}) \|_K (\| p - p_h \|_K + \| p_h - \tilde{p}_h \|_K)_K \\
\lesssim \sum_{K \in T_h} h_K \| \phi \|_{H^2(K)} (\| p - p_h \|_K + h_K \| S^{-1}u_h \|_K). \tag{3.25}
\]

From (3.24), it holds that

\[
\| p - p_h \|_K + h_K \| S^{-1}u_h \|_K \lesssim \mathcal{E}_K. \tag{3.26}
\]
According to (3.25) and (3.26), we arrive at
\[ I_4 \lesssim \sum_{K \in T_h} h_K \| \phi \|_{H^2(K)} \mathcal{E}_K, \]
which, together with the regularity estimate (3.3) of \( \phi \), implies the result (3.23).

**Lemma 3.5.** It holds that
\[ I_5 \lesssim e_h \| h \|_{L^\infty(\Omega)} \| Q_h p - p_h \|. \]  \hspace{1cm} (3.27)

**Proof.** From (3.26) it follows that
\[ I_5 = - \sum_{K \in T_h} (\nabla \cdot \mathbf{w}(\phi - \phi_h), p - p_h)_K. \]
\[ \lesssim \sum_{K \in T_h} \| \phi - \phi_h \|_K (\| p - p_h \|_K + \| p_h - \tilde{p}_h \|_K) \]
\[ \lesssim \sum_{K \in T_h} h_K \| \nabla \phi \|_K (\| p - p_h \|_K + h_K \| S^{-1} u_h \|_K) \]
\[ \lesssim e_h \| h \|_{L^\infty(\Omega)} \| \phi \|_{H^1}, \]
which, together with (3.2), yields the estimate (3.27).

**Lemma 3.6.** It holds that
\[ I_6 \lesssim e_h \| h \|_{L^\infty(\Omega)} \| Q_h p - p_h \|. \]  \hspace{1cm} (3.28)

**Proof.** For any \( \sigma \in \varepsilon_h^0 \), let \( \phi_\sigma \) denote the mean of \( \phi \) onto \( \sigma \), i.e.,
\[ \phi_\sigma := \frac{1}{|\sigma|} \int_\sigma \phi ds. \]

According to the continuity of the means of traces of the postprocessed scalar \( \tilde{p}_h \) (see Lemma 6.1 in [42]), and noticing \( \mathbf{w} \in RT_0(T_h) \), we have
\[ \int_\sigma [\tilde{p}_h] \mathbf{w} \cdot \mathbf{n} \phi = \int_\sigma [\tilde{p}_h] \mathbf{w} \cdot \mathbf{n} (\phi - \phi_\sigma) \lesssim \| [\tilde{p}_h] \|_\sigma \| \phi - \phi_\sigma \|_\sigma. \]  \hspace{1cm} (3.29)

A sidewise Poincaré inequality implies
\[ \| \phi - \phi_\sigma \|_\sigma \lesssim h_\sigma \| \gamma_\sigma (\nabla \phi) \|_\sigma \lesssim h_\sigma^{1/2} \| \phi \|_{H^2(\cup \omega_\sigma)}. \]  \hspace{1cm} (3.30)

In the last step above, we use trace inequality
\[ \| v \|_{\partial K} \lesssim h_K^{-1/2} \| v \|_K + \| v \|_K^{1/2} \| \nabla v \|_K^{1/2}, \quad \forall v \in H^1(K), \quad \forall K \in T_h, \]
and the local shape regularity of the mesh.

By the Bramble-Hilbert lemma and postprocessing technique, we have
\[ \| [\tilde{p}_h] \|_\sigma = \left\| \frac{1}{|\sigma|} \int_\sigma [\tilde{p}_h] ds \right\|_\sigma \lesssim h_\sigma \| [\gamma_\sigma (\nabla \tilde{p}_h)] \|_\sigma = h_\sigma \| [\gamma_\sigma (S^{-1} u_h)] \|_\sigma. \]  \hspace{1cm} (3.31)

A combination of (3.29)–(3.31) yields
\[ \int_\sigma [\tilde{p}_h] \mathbf{w} \cdot \mathbf{n} \phi \lesssim h_\sigma \| [\gamma_\sigma (S^{-1} u_h)] \|_\sigma h_\sigma^{1/2} \| \phi \|_{H^2(\cup \omega_\sigma)}. \]  \hspace{1cm} (3.32)

In light of the local shape regularity of elements, the above estimate leads to
\[ I_6 \lesssim \sum_{\sigma \in \varepsilon_h^0} h_\sigma \| [\gamma_\sigma (S^{-1} u_h)] \|_\sigma h_\sigma^{1/2} \| \phi \|_{H^2(\cup \omega_\sigma)}. \]
Recall a local efficiency result of $h^{1/2}|||\gamma_{\tau_h}(S^{-1}u_h)|||_{\sigma}$ as follows (see Lemma 7.2 in [21,25]),

$$h^{1/2}|||\gamma_{\tau_h}(S^{-1}u_h)|||_{\sigma} \lesssim ||S^{-1/2}(u - u_h)||_{h^r}.$$  

(3.34)

The desired result (3.28) follows from (3.33), (3.34) and (3.3).

□

We now give an estimate of $||Q_hp - p_h||$.

**Theorem 3.7.** Let $\epsilon_h$, osc$_h$, and $||h||_{L^\infty(\Omega)}$ denote the error of the stress and displacement variables given in (2.5), the oscillation of data given in (2.7), and the maximum norm of the mesh-size function, respectively, with respect to $T_h$. Then, under the weak regularity assumption (3.3) on $S$, there exits a positive constant $C_2$ only depending on the shape regularity of $T_h$, such that

$$||Q_hp - p_h|| \leq C_2C_D(||h||_{L^\infty(\Omega)}e_h + osc_h),$$  

(3.35)

where $C_D$ is one variant of coefficients.

**Proof.** The estimate (3.35) follows from a combination of (3.13), (3.14), (3.16), (3.21), (3.23), (3.27), and (3.28).

□

**Remark 3.8.** A combination of the two estimates (3.35) and (2.6) actually gives a posteriori bound for $Q_hp(p - p_h)$. Furthermore, following [9], we denote

$$L^k_{NC} := \left\{ q_h \in L^2(\Omega) : q_h|_K \in P_k(K), \forall K \in T_h, \sum_K \int_{\partial K} p_h \phi ds = 0, \forall \phi \in H^1_0(K) \right\},$$

and let $p_h^*$, $\hat{p}_h$ be respectively the interpolations in $L^k_{NC}$ of the interelement Lagrangian multiplier $\lambda_h$ and the displacement $p$ ([9, pp.186–187]; we note that in [9] $u$ represents the displacement variable and $p$ the stress variable). Following the same lines as in [9], the following estimate holds,

$$||\hat{p}_h - p_h^*|| \lesssim ||hS^{-1/2}(u - u_h)|| + ||Q_hp - p_h||,$$

which gives an a posteriori error estimate for $\hat{p}_h - p_h^*$.

**Remark 3.9.** For a pure diffusion problem, i.e., $w = r = 0$ in (1.1), it holds that $I_i = 0, i = 3, 4, 5, 6$. From the estimates of $I_1$ and $I_2$, we can obtain

$$||Q_hp - p_h|| \lesssim ||hS^{-1/2}(u - u_h)|| + osc_h,$$

which results in the quasi-orthogonality

$$(S^{-1}(u - u_h), u_h - u_H) \lesssim (||hS^{-1/2}(u - u_h)|| + osc_h)||f_h - f_H||,$$

where we have used the fact

$$\nabla \cdot u_h = Q_hz f = f_h \quad \text{and} \quad \nabla \cdot u_H = Q_hz f = f_H.$$

This estimate is somewhat different from the quasi-orthogonality results in [11,14,16,23,24].

**Remark 3.10.** If the mesh $T_h$ is quasi-uniform, then (3.35) is actually a superconvergence result of the displacement variable, i.e., $||Q_hp - p_h|| \leq C(u, p)h^2$. We refer to [2] on the superconvergence of mixed finite elements. Furthermore, if the mesh $T_h$ is uniform and regular, and the diffusion-dispersion $S$ is Lipschitz continuous, then, following the lines of the proof of Theorem 3.2 in [8], we obtain the superconvergence on the stress variable, $||S^{-1/2}(u_h - \Pi_H u)|| \leq C(u, p)h^{3/2}$. 

\[1336\]
4 Estimator reduction

Let $\omega_\sigma$ denote the patch of $\sigma \in \varepsilon_h$, and define $c_{\omega_\sigma}, D^2_{\omega_\sigma}, D^2_{\omega_\sigma}(K), D^2_{\omega_\sigma}$ respectively by

$$c_{\omega_\sigma} := \begin{cases} \max\left(\varepsilon_{\sigma,K}^{-1/2} c_{\sigma,K}^{-1/2}, \varepsilon_{\sigma,K}^{-1/2} \right), & \text{if } \sigma = K \cap L, \\ \varepsilon_{\sigma,K}^{-1/2}, & \text{if } \sigma \in \varepsilon_h \cap \partial \Omega, \end{cases} \quad D^2_{\omega_\sigma} := \max_{\sigma \in \varepsilon_h} D^2_{\sigma},$$

$$D^2_{\omega_\sigma}(K) := \max(h_K^2 D^2_{\sigma_1} c_{S,K}^{-1}, C_{DK}(\alpha_K, D^2_{\omega_\sigma})), \quad D^2_{\omega_\sigma} := \sum_{K \in T_h} D^2_{\omega_\sigma}(K),$$

where $C_{DK}$ in $D^2_{\omega_\sigma}(K)$ is given by

$$C_{DK} := 2 \max\left(\left(\frac{C_{S,K}^{1/2}}{\varepsilon_{S,K}} + \frac{h_K}{\sqrt{\varepsilon_{S,K}}} \|w\|_{L^\infty(K)}\right)^2, \frac{h_K^2 C_{S,K}^2}{c_{w,r,K}}\right).$$

**Lemma 4.1** (Estimator reduction). For a triangulation $T_H$ with $M_H \subset T_H$, let $T_h$ be a refinement of $T_H$ obtained by the AMFEM algorithm. Denote by $D^2_{\omega_\sigma}$ one variant of the coefficients onto the initial mesh $T_0$, and denote $\lambda := 1 - 2^{-b/2}$,

$$E^2_{\omega_\sigma} := \sum_{K \in T_h} \|S^{-1/2}(u_h - u_H)\|^2_K + c_{w,r,K} \|p_h - p_H\|^2_K. \quad (4.1)$$

Then for any $0 < \delta < 1$, it holds that

$$\eta^2_{\omega_\sigma}(u_h, p_h, T_h) \leq (1 + \delta)\{\eta^2_{\omega_\sigma}(u_H, p_H, T_H) - \lambda \eta^2_{\omega_\sigma}(u_H, p_H, M_H)\} + C^3_3(1 + \delta^{-1})D^2_{\omega_\sigma}E^2_{\omega_\sigma}. \quad (4.2)$$

**Proof.** By the triangle inequality and Young’s inequality, we have

$$h_K^2 \|S^{-1}u_h\|^2_K \leq (1 + \delta)h_K^2 \|S^{-1}u_H\|^2_K + (1 + \delta^{-1})h_K^2 c_{S,K}^{-1} \|S^{-1/2}(u_h - u_H)\|^2_K. \quad (4.3)$$

The inverse inequality implies $\|\nabla \cdot (u_h - u_H)\|_K \leq C_{\sigma,K}^{1/2}h_K^{-1} \|S^{-1/2}(u_h - u_H)\|_K$, which leads to

$$h_K^2 \|f - \nabla \cdot u_h + S^{-1}u_h \cdot w - (r + \nabla \cdot w)p_h\|^2_K \leq (1 + \delta)h_K^2 \|f - \nabla \cdot u_H + S^{-1}u_H \cdot w - (r + \nabla \cdot w)p_H\|^2_K + (1 + \delta^{-1})C_{DK} \|S^{-1/2}(u_h - u_H)\|^2_K + c_{w,r,K} \|p_h - p_H\|^2_K. \quad (4.4)$$

From the inverse inequality and local shape regularity of the mesh, it follows that

$$\|\gamma_{\omega_\sigma}(S^{-1}u_h)\|_\sigma \leq \|\gamma_{\omega_\sigma}(S^{-1}u_H)\|_\sigma + C_5c_{\omega_\sigma} h^{-1/2}_\sigma \|S^{-1/2}(u_h - u_H)\|_\sigma, \quad (4.5)$$

which, together with Young’s inequality, yields

$$h_\sigma \|\gamma_{\omega_\sigma}(S^{-1}u_h)\|^2_\sigma \leq (1 + \delta)h_\sigma \|\gamma_{\omega_\sigma}(S^{-1}u_H)\|^2_\sigma + (1 + \delta^{-1})C^2_{\sigma,K} c_{\omega_\sigma} \|S^{-1/2}(u_h - u_H)\|^2_\sigma. \quad (4.6)$$

Summing (4.3) and (4.4) over all the elements $K \in T_h$, summing (4.6) over all the interior sides $\sigma \in \varepsilon^0_h$, and noticing the monotonicity of variants of the coefficients, we obtain

$$\eta^2_{\omega_\sigma}(u_h, p_h, T_h) \leq (1 + \delta)\{\eta^2_{\omega_\sigma}(u_H, p_H, T_h) + C_3^3(1 + \delta^{-1})D^2_{\omega_\sigma}E^2_{\omega_\sigma}\}. \quad (4.7)$$

For a marked element $K \in M_H$, we set $T_{h,K} := \{K' \in T_h | K' \subset K\}$. It holds that

$$\left\{ \begin{array}{ll} \sum_{K' \in T_{h,K}} \eta^2_{\omega_\sigma}(u_H, p_H, K') \leq 2^{-b/2}\eta^2_{\omega_\sigma}(u_H, p_H, K), & \text{for } K \in T_H / T_h, \\ \eta^2_{\omega_\sigma}(u_H, p_H, K) \leq \eta^2_{\omega_\sigma}(u_H, p_H, K), & \text{for } K \in T_h / M_H, \end{array} \right.$$  

which results in the following estimate

$$\eta^2_{\omega_\sigma}(u_H, p_H, T_h) \leq 2^{-b/2}\eta^2_{\omega_\sigma}(u_H, p_H, M_H) + \eta^2_{\omega_\sigma}(u_H, p_H, T_h / T_h) = \eta^2_{\omega_\sigma}(u_H, p_H, T_h) - \lambda \eta^2_{\omega_\sigma}(u_H, p_H, M_H). \quad (4.8)$$

The desired result (4.2) follows from (4.7), (4.8) and the monotonicity of $D^2_{\omega_\sigma}$. \hfill \Box
5 Proof of Theorem 1.1

In this section, we show that the error plus some quantity uniformly reduces with a fixed factor on two successive meshes, which shows the AMFEM is convergent.

For convenience, we denote
\[ \zeta_h^2 := \| f - f_h \|^2 + \| h \nabla_h (S^{-1} u_h \cdot \mathbf{w}) \|^2, \]
where \( f_h \) is the \( L^2 \)-projection of \( f \) onto \( P_0(T_h) \). Note that \( \zeta_h \) is the same order of \( \eta_h \) and thus the term \( h_0^2 \zeta_h^2 \) is small if \( h_0 \) is sufficiently small.

**Lemma 5.1.** Let \((u_h, p_h) \in RT_0(T_h) \times P_0(T_h)\) be the approximation solutions to the stress and displacement variables with respect to \( T_h \), and \( e_h \) the error of the stress and displacement variables with respect to \( T_h \). Denote by \( h_0 \) the size of the quasi-uniform initial mesh \( T_0 \), by \( D_1 \) one variant of the coefficients determined by \( C_D \). Then it holds for any \( 0 < \delta_1 < 1 \) that
\[
\| S^{-1/2}(u - u_h) \|^2 \leq 2\delta_1^{-1} D_1 (h_0^2 e_h^2 + h_0^2 \zeta_h^2) + \| S^{-1/2}(u - u_H) \|^2 - \| S^{-1/2}(u_h - u_H) \|^2 + \frac{1}{2} \delta_1 \| \nabla \cdot (u_h - u_H) \|^2. \tag{5.1}
\]

**Proof.** From (2.1) and (2.3) we get
\[
\| S^{-1/2}(u - u_h) \|^2 = \| S^{-1/2}(u - u_H) \|^2 - \| S^{-1/2}(u_h - u_H) \|^2 - 2(S^{-1}(u - u_h), u_h - u_H) 
\]
\[ = \| S^{-1/2}(u - u_H) \|^2 - \| S^{-1/2}(u_h - u_H) \|^2 - 2(p - p_h, \nabla \cdot (u_h - u_H)) = \| S^{-1/2}(u - u_H) \|^2 - \| S^{-1/2}(u_h - u_H) \|^2 - 2(Q_h p - p_h, \nabla \cdot (u_h - u_H)). \tag{5.2}
\]

The definition of the residual \( R_K \) and the assumptions of data imply that for each \( K \in T_h \),
\[ h_K \| R_K - \tilde{R}_K \|_K \leq C_0 (h_0^2 \| \nabla (S^{-1} u_h \cdot \mathbf{w}) \|_K + h_K \| f - f_h \|_K), \]
which, together with the fact \( h \| \nabla u \|_{L^\infty(\Omega)} \leq h_0 \) and the definition (2.7) of \( \text{osc}_h \), yields the estimate
\[ \text{osc}_h^2 \leq 2C_0^2 h_0^2 \zeta_h^2. \tag{5.3}
\]

Applying the above estimate (5.3) to (3.35), we obtain
\[ \| Q_h p - p_h \|^2 \leq D_1 (h_0^2 e_h^2 + h_0^2 \zeta_h^2). \tag{5.4}
\]
In light of Young’s inequality, we have
\[ 2(Q_h p - p_h, \nabla \cdot (u_h - u_H)) \leq 2\delta_1^{-1} \| Q_h p - p_h \|^2 + \frac{\delta_1}{2} \| \nabla \cdot (u_h - u_H) \|^2. \tag{5.5}
\]
The desired result (5.1) follows from a combination of (5.2), (5.4), and (5.5). \qed

**Lemma 5.2.** Let \( D_2 \) and \( D_3 \) be two variants of the coefficients respectively given by
\[ D_2 := \max_{K \in T_h} \| w \|_{L^\infty(K)}^{-1}, \quad D_3 := \max_{K \in T_h} C_{w, r, K}. \]
Under the assumptions of Lemma 5.1, it holds that
\[
\| \nabla \cdot (u - u_h) \|^2 \leq 4D_1 D_3 (h_0^2 e_h^2 + h_0^2 \zeta_h^2) + \| \nabla \cdot (u - u_H) \|^2 - \frac{1}{2} \| \nabla \cdot (u_h - u_H) \|^2 + 4D_2 \| S^{-1/2}(u - u_h) \|^2. \tag{5.6}
\]
Proof. Notice
\[
\| \nabla \cdot (u - u_h) \|^2 = \| \nabla \cdot (u - u_H) \|^2 - \| \nabla \cdot (u_h - u_H) \|^2 - 2(\nabla \cdot (u - u_h), \nabla \cdot (u_h - u_H)).
\] (5.7)
A combination of (2.2) and (2.4) leads to
\[
(\nabla \cdot (u - u_h), \nabla \cdot (u_h - u_H)) = (S^{-1}(u - u_h) \cdot w, \nabla \cdot (u_h - u_H)) \\
- ((r + \nabla \cdot w)(Q_h p - p_h), \nabla \cdot (u_h - u_H)) \\
\leq 2D_2 \| S^{-1}(u - u_h) \|^2 + 2D_3 \| Q_h p - p_h \|^2 \\
+ \frac{1}{4} \| \nabla \cdot (u_h - u_H) \|^2.
\] (5.8)
The estimate (5.6) follows from (5.7), (5.4) and (5.8).

Lemma 5.3. Let \( D_4 \) be one variant of the coefficients given by \( D_4 := \max_{K \in T_h} c_{w,r,K} \). Under the assumption of Lemma 5.1, it holds that
\[
\sum_{K \in T_h} c_{w,r,K} \| p - p_h \|^2 \leq \sum_{K \in T_h} c_{w,r,K} \| p - p_H \|^2 - \frac{1}{2} \sum_{K \in T_h} c_{w,r,K} \| p_h - p_H \|^2 \\
+ 2D_4 D_1 \left( h_0^2 \zeta_h^2 + h_0^2 \zeta_h^2 \right). 
\] (5.9)
Proof. For each element \( K \in T_h \), the following identity holds,
\[
\| p - p_h \|^2 \leq \| p - p_H \|^2 - \| p_h - p_H \|^2 - 2(p - p_h, p_h - p_H)_K \\
= \| p - p_H \|^2 - \| p_h - p_H \|^2 - 2(Q_h p - p_h, p_h - p_H)_K.
\] (5.10)
Notice that \( c_{w,r,K} \) does not change from \( T_h \) to \( T_h \). Summing (5.10) by multiplying \( c_{w,r,K} \) over all the elements \( K \in T_h \), we have
\[
\sum_{K \in T_h} c_{w,r,K} \| p - p_h \|^2 \leq \sum_{K \in T_h} c_{w,r,K} \| p - p_H \|^2 - \frac{1}{2} \sum_{K \in T_h} c_{w,r,K} \| p_h - p_H \|^2 \\
+ 2D_4 \| Q_h p - p_h \|^2.
\] (5.11)
The conclusion (5.9) follows from (5.11) and (5.4).

In what follows, we show the reduction of the error. To this end, set \( \gamma_1, \gamma_2, \varepsilon_0 \), and \( \delta_1 \) to be positive constants, which will be determined below. Introduce the following quantity:
\[
A_h^2 := \delta_1 (1 - \varepsilon_0)^{-1} \| \nabla \cdot (u - u_h) \|^2 + \gamma_1 h_0^2 + \gamma_2 \zeta_h^2.
\] (5.12)
Note that the definition of \( A_h \) is similar to \( A_h \).

Theorem 5.4. Let \((u_h, p_h) \in RT_0(T_h) \times P_0(T_h)\) and \((u_H, p_H) \in RT_0(T_H) \times P_0(T_H)\) be the approximated solutions to the stress and displacement variables with respect to \( T_h \) and \( T_H \), respectively. Denote by \( e_h \) and \( e_H \) the errors of the stress and displacement variables with respect to \( T_h \) and \( T_H \), respectively. Let \( h_0 \) be the size of the quasi-uniform initial mesh \( T_0 \), and \( q \) and \( \alpha \in (0, 1) \) two constants to be determined below. Then, under the weak regularity assumption (3.3) on \( S \), it holds that
\[
e_h^2 + (1 - h_0 q) A_h^2 \leq \alpha^2 \left( e_H^2 + (1 - h_0 q) A_H^2 \right),
\] (5.13)
when \( h_0 \leq \frac{1-\alpha^2}{1+\alpha^2} q \).
Proof. For convenience, denote
\[
D_5 := 2D_1 + 4\delta_1^2 D_3 + 2\delta_1 D_1 D_4.
\]
We next choose $1340\varepsilon^2 \delta \eta_h$. From the definition of $\delta$, for any given $D_w \leq h^2$. The reliable estimate, (2.6), of the stress and displacement variables, together with (5.14), implies

$$\varepsilon^2 + \delta_1 \| \nabla \cdot (u - u_h) \|^2 \leq \varepsilon^2 + \delta_1 \| \nabla \cdot (u - u_H) \|^2 - \frac{1}{2} E_H^2 + C_1 D_5 \delta_1^{-1} h_0^2 \eta_h^2 + \varepsilon \varepsilon + D_5 \delta_1^{-1} h_0^2 \eta_h^2,$$

which results in the following inequality:

$$\varepsilon^2 + \frac{\delta_1}{\varepsilon_0} \| \nabla \cdot (u - u_h) \|^2 \leq (1 + \varepsilon) \varepsilon^2 + \frac{\delta_1}{\varepsilon_0} \| \nabla \cdot (u - u_H) \|^2 - \frac{1}{2} E_H^2 + \frac{C_1 D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2 + \frac{D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2.$$

According to the triangle inequality and the inverse inequality, it holds for each $K \in T_h$ that

$$h_K \| \nabla (S^{-1} u_h \cdot w) \|_K \leq h_K \| \nabla (S^{-1} u_H \cdot w) \|_K + C_7 \| S^{-1} (u_h - u_H) \cdot w \|_K.$$

Notice that $\| f - f_h \|_K \leq \| f - f_H \|_K$ for all $K \in T_h$. For any given $\delta_3 > 0$ which will be determined below, (5.16) and Young’s inequality imply

$$\zeta_h^2 \leq (1 + \delta_3) \varepsilon^2 + (1 + \delta_3^{-1}) C_7 D_2 \| S^{-1/2} (u_h - u_H) \|^2.$$

From the definition of $A_h^2 (5.12)$, the estimator reduction (4.2) with the marking strategy, the estimates (5.15) and (5.17), and the fact $\| S^{-1/2} (u_h - u_H) \|^2 \leq E_H^2$, for any given $\delta_2 > 0$ which will be determined below, it holds that

$$e_h^2 + A_h^2 \leq (1 + \varepsilon) \varepsilon^2 + \frac{\delta_1}{\varepsilon_0} \| \nabla \cdot (u - u_H) \|^2 - \frac{1}{2} E_H^2 + \frac{C_1 D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2 + C_1 \gamma_1 E_H^2 + \gamma_1 (1 + \delta_2)(1 - \lambda \theta^2) \eta_h^2 + \gamma_2 C_7^2 D_2 \left(1 + \frac{1}{\delta_3}\right) E_H^2 + \frac{D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2 + \gamma_2 (1 + \delta_3) \zeta_h^2.$$

We next choose $\gamma_1$ and $\gamma_2$ such that

$$\gamma_1 C_7^2 (1 + \delta_2^{-1}) D_2^2 = \frac{1}{4(1 - \varepsilon_0)}, \quad \gamma_2 C_7^2 D_2 (1 + \delta_3^{-1}) = \frac{1}{4(1 - \varepsilon_0)}.$$

Then it follows that

$$e_h^2 + A_h^2 \leq (1 + \varepsilon) \varepsilon^2 + \frac{\delta_1}{\varepsilon_0} \| \nabla \cdot (u - u_H) \|^2 + \gamma_1 (1 + \delta_2)(1 - \lambda \theta^2) \eta_h^2 + \frac{D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2 + \gamma_2 (1 + \delta_3) \zeta_h^2.$$

For any given $\delta_4, \delta_5 > 0$ which will be determined below, the reliable estimate (2.6) on $T_H$, i.e., $e_h^2 \leq C_4 \eta_h^2$, and the above estimate (5.19), indicate

$$e_h^2 + A_h^2 \leq \left(1 + \varepsilon - \frac{1}{2} \lambda \theta^2 \gamma_1 (1 + \delta_2) C_4^{-1}\right) e_h^2 + \gamma_1 (1 + \delta_2) \left(1 - \frac{1}{2} \lambda \theta^2\right) \eta_h^2 + \delta_4 (1 - \varepsilon_0) \| \nabla \cdot (u - u_H) \|^2 + \frac{D_5}{\delta_1(1 - \varepsilon_0)} h_0^2 \eta_h^2.$$
We finally choose \( \text{Remark 5.5} \) (Choices of the initial mesh size) which implies

These choices, together with (5.20) and (5.21), lead to

From (5.22) we arrive at

which yields the assertion (5.13) with

In fact, we may firstly fix \( \delta_2 \) and, subsequently, a sufficiently small \( \varepsilon \) such that

Let \( D_6 \) be one variant of the coefficients given by

From (3.19), we get

We further choose \( \delta_i \) \((i = 3, 4, 5)\) such that

In fact, we may firstly fix \( \delta_3 \) satisfying \( \delta_3 < C_8 h_0 \), then choose \( \delta_5 \) such that

Finally, by noticing the choice of \( \gamma_2 \), we can choose \( \delta_4 \) with

These choices, together with (5.20) and (5.21), lead to

Let \( q \) be one variant of the coefficients given by

From (5.22) we arrive at

which implies

We finally choose \( h_0 \) such that

which yields the assertion (5.13) with \( h_0 \leq (1 - \alpha^2)/(q(1 + \alpha^2)) \).

\[ \square \]

**Remark 5.5** (Choices of the initial mesh size). Some simple calculations show

\[
q \leq \max \left\{ D(\delta_1, \delta_2), \frac{4h_0C_2^2D_2D_5}{\delta_1} (1 + \delta_3^{-1}) \right\}
\]
with
\[ D(\delta_1, \delta_2) := \max \left\{ C_8 + 1, \frac{C_2^2 (1 + \delta_2^{-1}) D^2_{T_0}}{C_2^2 D}, \frac{4 C_2^2 C_1 D_5 (1 + \delta_2^{-1}) D^2_{T_0}}{\delta_1} \right\}. \]

Then it holds that
\[ \frac{1}{q} \geq \min \left\{ \frac{1}{D(\delta_1, \delta_2)}, \frac{\delta_1}{4 h_0 C_2^2 D_2 D_5 (1 + \delta_3^{-1})} \right\}, \]
which indicates
\[ \frac{1}{q} \geq \frac{1}{D(\delta_1, \delta_2)} \text{ if } h_0 \leq \frac{\delta_1}{4 C_2^2 D_2 D_5 (1 + \delta_3^{-1})}. \]

As required in Theorem 5.4, the initial mesh size \( h_0 \) is assumed to satisfy \( h_0 \leq \frac{1 - \alpha^2}{\alpha^2 (1 + \alpha^2)} \). Then eventually we may choose \( h_0 \) with
\[ h_0 \leq \min \left\{ \frac{1 - \alpha^2}{(1 + \alpha^2) D(\delta_1, \delta_2)}, \frac{\delta_1}{4 C_2^2 D_2 D_5 (1 + \delta_3^{-1})} \right\}. \]

**Proof of Theorem 1.1.** Theorem 5.4 shows that the error of the stress and displacement variables plus the quantity \( A_2^h \) uniformly reduces with a fixed factor \( \alpha^2 \) between two successive meshes. Replace the subscripts \( H \) and \( h \) respectively by the iteration counters \( k \) and \( k+1 \), we then obtain Theorem 1.1 directly from Theorem 5.4. \( \square \)

### 6 Numerical experiments

In this section, we test the performance of the adaptive algorithm AMFEM described in Section 2 with four model problems. We are thus able to study how meshes adapt to the various effects from lack of regularity of solutions and convexity of domains to data smoothness, boundary layers and changing boundary conditions. We note that the implementation of AMFEM is done without enforcing the interior node property in the refinement step.

#### 6.1 Model problem with singularity at the origin

We consider the problem (1.1) in an \( L \)-shape domain \( \Omega = \{(−1,1) \times (0,1)\} \cup \{(−1,0) \times (−1,0)\} \) with \( w = r = 0 \) and \( f = 0 \). The exact solution is given by
\[ p(\rho, \theta) = \rho^2 \sin \left( \frac{2 \theta}{3} \right), \]
where \( \rho, \theta \) are the polar coordinates.

Since this model possesses singularity at the origin, we see in Figure 1(a) that the refinement concentrates around the origin, which means the predicted error estimator captures well the singularity of the solution. Figure 1(b) reports the estimated and actual errors of the numerical solutions on uniformly and adaptively refined meshes. It can be seen that the error of the stress and displacement in \( L^2 \) norm uniformly reduces with a fixed factor on two successive meshes after several steps of iterations, and that the error on the adaptively refined meshes decreases more rapidly than the one on the uniformly refined meshes. This shows that the adaptive mixed finite element method is convergent with respect to the energy error.

#### 6.2 Model problem with inhomogeneous diffusion tensor

We consider the problem (1.1) in a square domain \( \Omega = (−1,1) \times (−1,1) \) with \( w = r = 0 \) and \( f = 0 \), where \( \Omega \) is divided into four subdomains \( \Omega_i \ (i = 1, 2, 3, 4) \) corresponding to the axis quadrants (in the counterclockwise direction), and the diffusion-dispersion tensor \( S \) is piecewise constant with \( S = s_i I \) in
Figure 1 A mesh with 14692 triangles (a) and the estimated and actual errors in uniformly/adaptively refined meshes (b) for the marking parameter $\theta = 0.5$.

$\Omega_i$. This model problem is taken from [26, 39, 42]. We suppose the exact solution of this model has the form

$$p(\rho, \theta) = \rho^\alpha (a_i \sin(\alpha \theta) + b_i \cos(\alpha \theta))$$

in each $\Omega_i$ with the Dirichlet boundary conditions. Here $\rho, \theta$ are the polar coordinates in $\Omega$, $a_i$ and $b_i$ are constants depending on $\Omega_i$, and $\alpha$ is a parameter. We note that the stress solution, $u = -s\nabla P$, is not continuous across the interfaces, and only its normal component is continuous. It finally exhibits a strong singularity at the origin. We consider two sets of coefficients in Table 1.

In the MARK step, the marking parameter $\theta$, in terms of Dörfler marking, is chosen as 0.7 in the first case and as 0.94 in the second case. Table 2 shows for Case 1 some results of the actual error $e_k$, the a posteriori indicator $\eta_k$, the experimental convergence rate, $\text{EOC}_E$, of $E_k$, and the experimental convergence rate, $\text{EOC}_\eta$, of $\eta_k$, where

$$\text{EOC}_E := \frac{\log(e_k / e_{k-1})}{\log(\text{DOF}_k / \text{DOF}_{k-1})}, \quad \text{EOC}_\eta := \frac{\log(\eta_k / \eta_{k-1})}{\log(\text{DOF}_k / \text{DOF}_{k-1})},$$

| $k$ | DOF$_k$ | $e_k$ | $\eta_k$ | $\text{EOC}_E$ | $\text{EOC}_\eta$ |
|-----|---------|-------|----------|----------------|----------------|
| 1   | 8       | 1.3665 | 5.0938   | –              | –              |
| 2   | 20      | 1.1346 | 3.4700   | 0.2030         | 0.4189         |
| 9   | 2235    | 0.1776 | 1.1115   | 0.4016         | 0.4004         |
| 11  | 7165    | 0.1106 | 0.7111   | 0.3851         | 0.4004         |
| 12  | 13188   | 0.0871 | 0.5566   | 0.3915         | 0.4015         |
| 14  | 43785   | 0.0510 | 0.3365   | 0.4707         | 0.4476         |
| 15  | 76770   | 0.0387 | 0.2581   | 0.4915         | 0.4724         |

Table 1 Two sets of coefficients

| Case 1 | Case 2 |
|--------|--------|
| $s_1 = s_3 = 5, s_2 = s_4 = 1, \alpha = 0.53544095$ | $s_1 = s_3 = 100, s_2 = s_4 = 1, \alpha = 0.12690207$ |
| $a_1 = 0.44721360, b_1 = 1.00000000$ | $a_1 = 0.10000000, b_1 = 1.00000000$ |
| $a_2 = -0.74535599, b_2 = 2.33333333$ | $a_2 = -9.60396040, b_2 = 2.96039604$ |
| $a_3 = -0.94411759, b_3 = 0.55555555$ | $a_3 = -0.48035487, b_3 = -0.88275659$ |
| $a_4 = -2.40170264, b_4 = -0.81481418$ | $a_4 = 7.70156488, b_4 = -6.45646175$ |
and $\text{DOF}_k$ denotes the number of elements with respect to the $k$-th iteration. We can see that the convergence rates $\text{EOC}_E$ and $\text{EOC}_\eta$ are close to 0.5 as $k = 15$, which means the optimal decays of the actual error $e_k$ and the a posteriori error indicator $\eta_k$ are almost attained after 15 iterations with optimal meshes.

Figure 2 shows an adaptively refined mesh with 4763 elements and the estimated and actual errors against the number of elements in adaptively refined meshes for Case 1. Figure 3 shows an adaptively refined mesh with 1093 elements and the actual error against the number of elements in adaptively refined meshes for Case 2.

From Figures 2(a) and 3(a), we can see that the refinement concentrates around the origin, which means the AMFEM algorithm detects the region of rapid variation. In Figures 2(b) and 3(b) each includes an optimal convergence line, which shows in both cases, the energy error performs a trend of descending with an optimal order convergence rate after several steps of adaptive iterations for the problem with strongly discontinuous coefficients. We note that the energy error is approximated with a 7-point quadrature formula in each triangle.

In the following two examples, we consider the convection-dominated case with $S = \varepsilon I$.

![Figure 2](image1.png)  
**Figure 2** A mesh with 4763 triangles (a) and the estimated and actual errors against the number of elements in adaptively refined meshes (b): Case 1

![Figure 3](image2.png)  
**Figure 3** A mesh with 1093 triangles (a) and the actual error against the number of elements in adaptively refined mesh (b): Case 2
6.3 Convection-dominated model problem: boundary layers

In this example, we take $\Omega = (0,1) \times (0,1)$ in $\mathbb{R}^2$, and choose $w = (1,1)$ and $r = 0$. Furthermore, we set $p = 0$ on $\partial \Omega$, and select the right-hand side $f$ such that the analytical solution to (1.1) is given by

$$p(x,y) = \left( \frac{\exp(\frac{x-1}{\varepsilon}) - 1}{\exp(-\frac{1}{\varepsilon}) - 1} + x - 1 \right) \left( \frac{\exp(\frac{y-1}{\varepsilon}) - 1}{\exp(-\frac{1}{\varepsilon}) - 1} + y - 1 \right).$$

The solution is smooth, but has boundary layers at $x = 1$ and $y = 1$, with the layer width of order $O(\varepsilon)$. This problem is well-suited to test whether the estimator is able to pick up the steep gradients near these boundaries.

We start computations from the original mesh consisting of 8 right-angled triangles, and we choose the marking parameter $\theta = 0.5$ in the adaptive algorithm AMFEM.

Figure 4 shows the mesh with 10838 triangles (a) and the postprocessing approximation to the scalar displacement $p$ on the corresponding adaptively refined mesh (b) in the case $\varepsilon = 0.01$. Here the value of the postprocessing approximation on each vertex is taken as the algorithmic mean of the values of the displacement finite element solution on all the elements sharing the vertex. The reason for the postprocessing is that the displacement finite element solution is not continuous on each vertex of the triangulation. We see that the refinement focuses around boundary layers, which indicates that the estimators actually capture boundary layers and resolve them in convection-dominated regions. In addition, the postprocessing approximation to the scalar displacement obtains satisfactory results.

Figure 5 shows the actual error (energy error) results against the number of elements in adaptively refined meshes for $\varepsilon = 0.1$ (a) and $\varepsilon = 0.01$ (b), including two theoretically-optimal order $(-1/2)$ convergence lines. We see that in each case the actual error descends almost at the optimal rate of convergence after several steps of iterations. The numerical results confirm our theoretical analysis.

6.4 Convection-dominated model problem: Interior and boundary layers

Set the domain $\Omega = [-1,1] \times [-1,1]$ with the non-homogeneous Dirichlet boundary conditions, the velocity field $w = (2,1)$, and the reaction term $r = 0$ in (1.1). The source term $f = 0$, the Dirichlet boundary conditions are as follows: $p = 0$ along the left and top sides of the square and $p = 100$ along the right and bottom sides. The exact solution of this problem is unknown, but it is known that it exhibits an exponential boundary layer at the boundary $x = 1, y > 0$ and a parabolic interior layer along the line connecting the points $(-1,-1)$ and $(1,0)$.
Figure 5  Actual error against the number of elements in adaptively refined meshes for $\varepsilon = 0.1$ (a) and $\varepsilon = 0.01$ (b) for the marking parameter $\theta = 0.5$.

Figure 6  A mesh with 47324 elements (a) for $\varepsilon = 0.01, \theta = 0.8$ and estimated error against the number of elements in adaptively refined meshes (b) for $\varepsilon = 0.1, \theta = 0.5$.

We still perform the AMFEM algorithm described in Section 2 from the original mesh consisting of 8 right-angled triangles. From Figure 6(a), we can see that when using adaptive refinement the mesh concentrates close to the exponential and parabolic layers. We note that the refinement first occurs close to the region $x = 1, y > 0$, since the exponential layer is stronger than the parabolic layer. Figure 6(a) also illustrates that the a posteriori error estimator exactly captures the behavior of the solution. Figure 6(b) shows that the estimated error rapidly reduces starting from the fourth step of iterations, and reaches the optimal rate $(−1/2)$ of convergence until the seventeenth step. This convergence result is consistent with our theoretical analysis.

Acknowledgements  This work was supported by Education Science Foundation of Chongqing (Grant No. KJ120420), National Natural Science Foundation of China (Grant No. 11171239), Major Research Plan of National Natural Science Foundation of China (Grant No. 91430105) and Open Fund of Key Laboratory of Mountain Hazards and Earth Surface Processes, Chinese Academy Sciences.

References

1 Ainsworth M, Oden J T. A Posteriori Error Estimation in Finite Element Analysis. New York: John Wiley & Sons, 2000
2 Arnold D N, Brezzi F. Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates. Math Model Numer Anal, 1985, 19: 7–32
3 Arnold D N, Falk R S, Winther R. Finite element exterior calculus, homological techniques, and applications. Acta Numer, 2006, 15: 1–155
4 Babuška I, Rheinboldt W. A-posteriori error estimates for the finite element method. Internat J Numer Methods Engrg, 1978, 12: 1597–1615
5 Babuška I, Rheinboldt W. Error estimates for adaptive finite element computations. SIAM J Numer Anal, 1978, 15: 736–754
6 Babuška I, Strouboulis T. The Finite Element Method and Its Reliability. Oxford: Clarendon Press, 2001
7 Becker R, Mao S P. An optimally convergent adaptive mixed finite element method. Numer Math, 2008, 111: 35–54
8 Brandts J. Superconvergence and a posteriori error estimation for triangular mixed finite elements. Numer Math, 1994, 68: 311–324
9 Brezzi F, Fortin M. Mixed and Hybrid Finite Element Methods. New York: Springer-Verlag, 1991
10 Carstensen C. A posteriori error estimate for the mixed finite element method. Math Comp, 1977, 66: 465–476
11 Carstensen C, Hoppe R H W. Error reduction and convergence for an adaptive mixed finite element method. Math Comp, 2006, 75: 1033–1042
12 Carstensen C, Hu J. A unifying theory of a posteriori error control for nonconforming finite element methods. Numer Math, 2007, 107: 473–502
13 Carstensen C, Hu J, Orlando A. Framework for the a posteriori error analysis of nonconforming finite elements. SIAM J Numer Anal, 2007, 45: 68–82
14 Carstensen C, Rabus H. An optimal adaptive mixed finite element method. Math Comp, 2011, 80: 649–667
15 Cascon J M, Kreuzer C, Nochetto R H, et al. Quasi-optimal convergence rate for an adaptive finite element method. SIAM J Numer Anal, 2008, 46: 2524–2550
16 Chen L, Holst M, Xu J C. Convergence and optimality of adaptive mixed finite element methods. Math Comp, 2009, 78: 35–53
17 Ciarlet P G. The Finite Element Method for Elliptic Problems. Amsterdam-New York-Oxford: North-Holland, 1978
18 Demlow A, Stevenson R. Convergence and quasi-optimality of a adaptive finite element method for controlling $L^2$ errors. Numer Math, 2011, 117: 185–218
19 Dörfler W. A convergent adaptive algorithm for Poisson’s equation. SIAM J Numer Anal, 1996, 33: 1106–1124
20 Douglas J R, Roberts J E. Global estimates for mixed methods for second order elliptic equations. Math Comp, 1985, 44: 39–52
21 Du S H. A new residual posteriori error estimates of mixed finite element methods for convection-diffusion-reaction equations. Numer Methods Partial Differential Equations, 2014, 30: 593–624
22 Du S H, Xie X P. Residual-based a posteriori error estimates of nonconforming finite element method for elliptic problems with Dirac delta source terms. Sci China Ser A, 2008, 51: 1440–1460
23 Du S H, Xie X P. Error reduction, convergence and optimality of an adaptive mixed finite element method. J Syst Sci Complex, 2012, 25: 195–208
24 Du S H, Xie X P. Error reduction, convergence and optimality for adaptive mixed finite element methods for diffusion equations. J Comput Math, 2012, 30: 483–503
25 Du S H, Xie X P. On residual-based a posteriori error estimators for lowest-order Raviart-Thomas element approximation to convection-diffusion-reaction equations. J Comput Math, 2014, 32: 522–546
26 Eigestad G T, Klausen R A. On the convergence of the multi-point flux approximation $O$-method: Numerical experiments for discontinuous permeability. Numer Methods Partial Differential Equations, 2005, 21: 1079–1098
27 Hiptmair R. Canonical construction of finite elements. Math Comp, 1999, 68: 1325–1346
28 Hu J, Shi Z C, Xu J C. Convergence and optimality of the adaptive Morley element method. Numer Math, 2012, 121: 731–752
29 Jia S H, Chen H T, Xie H H. A posteriori error estimator for eigenvalue problems by mixed finite element method. Sci China Math, 2013, 56: 887–900
30 Mekchay K, Nochetto R H. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. SIAM J Numer Anal, 2005, 43: 1803–1827
31 Mitchell W F. A comparison of adaptive refinement techniques for elliptic problems. ACM Trans Math Software, 1989, 15: 326–347
32 Mitchell W F. Optimal multilevel iterative methods for adaptive grids. SIAM J Sci Comput, 1992, 13: 146–167
33 Morin P, Nochetto R H, Siebert K G. Data oscillation and convergence of adaptive FEM. SIAM J Numer Anal, 2000, 38: 466–488
34 Morin P, Nochetto R H, Siebert K G. Convergence of adaptive finite element methods. SIAM Rev, 2002, 44: 631–658
35 Morin P, Nochetto R H, Siebert K G. Local problems on stars: A posteriori error estimators, convergence, and performance. Math Comp, 2003, 72: 1067–1097
36 Raviart P A, Thomas J M. A mixed finite element method for 2nd order elliptic problems. In: Mathematical Aspects of Finite Element Methods. Lecture Notes in Mathematics, vol. 606. Berlin: Springer, 1977, 292–315
37 Rivara M C. Mesh refinement processes based on the generalized bisection of simplices. SIAM J Numer Anal, 1984, 21: 604–613
38 Rivara M C. Design and data structure for fully adaptive, multigrid finite-element software. ACM Trans Math Software, 1984, 10: 242–264
39 Rivière B, Wheeler M F. A posteriori error estimates for a discontinuous Galerkin method applied to elliptic problems. Log number: R74. Comput Math Appl, 2003, 46: 141–163
40 Sewell E G. Automatic generation of triangulations for piecewise polynomial approximation. PhD dissertation. West Lafayette: Purdue University, 1972
41 Verfürth R. A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques. Chichester-New York: Wiley-Teubner, 1996
42 Vohralík M. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. SIAM J Numer Anal, 2007, 45: 1570–1599
43 Yang Y D, Bi H. Local a priori/a posteriori error estimates of conforming finite elements approximation for Steklov eigenvalue problems. Sci China Math, 2014, 57: 1319–1329
44 Yang Y D, Jiang W. Upper spectral bounds and a posteriori error analysis of several mixed finite element approximations for the Stokes eigenvalue problem. Sci China Math, 2013, 56: 1313–1330