CONVERGENCE AND OPTIMALITY OF ADAPTIVE MIXED FINITE ELEMENT METHODS

LONG CHEN, MICHAEL HOLST, AND JINCHAO XU

Abstract. The convergence and optimality of adaptive mixed finite element methods for the Poisson equation are established in this paper. The main difficulty for mixed finite element methods is the lack of minimization principle and thus the failure of orthogonality. A quasi-orthogonality property is proved using the fact that the error is orthogonal to the divergence free subspace, while the part of the error containing divergence can be bounded by the data oscillation using a discrete stability result. This discrete stability result is also used to get a localized discrete upper bound which is crucial for the proof of the optimality of the adaptive approximation.

1. INTRODUCTION

Adaptive methods are now widely used in scientific computation to achieve better accuracy with minimum degrees of freedom. While these methods have been shown to be very successful, the theory ensuring the convergence of the algorithm and the advantages over non-adaptive methods is still under development. Recently, several results have been obtained for standard finite element methods for elliptic partial differential equations [7, 34, 44, 46, 11, 55, 47, 26].

In this paper, we shall establish the convergence and optimality of adaptive mixed finite element methods (AMFEMs) of the model problem

\[
-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial \Omega,
\]

posed on a polygonal and simply connected domain \( \Omega \subset \mathbb{R}^2 \). In many applications ([23]) the variable \( \sigma = -\nabla u \) is of interest and it is therefore convenient to use mixed finite element methods, such as the Raviart-Thomas mixed method [49] and Brezzi-Douglas-Marini mixed method [22]. We shall construct adaptive mixed finite element methods based on the local refinement of triangulations and prove they will produce a sequence of approximation of \( \sigma \) in an optimal way.

Our main result is the following optimal convergence of our algorithms AMFEM and its variant. Let \( \sigma_N \) be the approximation of \( \sigma \) based on the triangulation \( T_N \) obtained in AMFEM. If \( \sigma \in A^\ast \) and \( f \in A_s^\ast \), then

\[
\| \sigma - \sigma_N \| \leq C(\| \sigma \|_{A^\ast} + \| f \|_{A_s^\ast})(\# T_N - \# T_0)^{-s},
\]
where \((\mathcal{A}^s, \| \cdot \|_{\mathcal{A}^s})\) and \((\mathcal{A}_0^s, \| \cdot \|_{\mathcal{A}_0^s})\) are approximation spaces as in [11]. The index \(s\) is used to characterize the best possible approximation rate of \(\sigma\), which depends on the regularity of the solution and data, and the order of elements. For example when \(f \in L^2(\Omega)\) and \(\sigma \in W^{1,1}\), we can achieve the optimal convergence rate \(s = 1/2\) for the lowest order Raviart-Thomas finite element space. We refer to [12] for the characterization of \(\mathcal{A}^s\) in terms of Besov spaces and to [8, 9, 33, 32] for the regularity results in Besov norms. We comment that to apply our adaptive algorithm, we do not need to know \(s\) explicitly. Our algorithm will produce the best possible approximation rate for the unknown \(\sigma\).

For the analysis of the convergence of adaptive procedure, we follow the new approach by Cascon, Kreuzer, Nochetto and Siebert [26], and for the optimality we mainly use the simplified case in Stevenson’s work [55]. A distinguish feature of the new approach for the convergence proof is the relaxation of the interior node requirement for the refinement. We do not claim any originality on the proof of convergence and optimality. Instead the main contribution of this paper is to establish two important ingredients used in the proof, namely quasi-orthogonality and discrete upper bound.

One main ingredient in the convergence analysis of standard AFEM is that the error is orthogonal to the finite element spaces in energy-related inner product since the standard finite element approximation can be characterized as a minimizer of Dirichlet-type energy. For mixed finite element methods, however, the approximation is a saddle point of the corresponding energy and thus there is no orthogonality available. We shall prove a quasi-orthogonality result. A similar result for the lowest order Raviart-Thomas finite element space has recently been proved by Carstensen and Hoppe [25], where a special relation between mixed finite element method and non-conforming method is used. In this paper, we shall propose a new and more straight-forward approach which works for any order elements and both Raviart-Thomas and Brezzi-Douglas-Marini methods. The main observation is that the error is orthogonal to the divergence free subspace, while the part of the error containing divergence can be bounded by the data oscillation using a discrete stability result.

Another ingredient to establish the optimality of the adaptive algorithm is the localized discrete upper bound for \textit{a posteriori} error estimator. Using the discrete stability result, we are able to obtain such discrete upper bound and use it to prove the optimality of the convergent algorithm. The optimality of mixed adaptive finite element methods seems to be new.

The rest of this paper is organized as follows. In Section 2, we shall introduce mixed finite element methods and give a short review of mesh adaptivity through local refinement. We shall include many preliminary results in this section for later usage. In Section 3, we shall prove the discrete stability result and use it to prove the quasi-orthogonality result. In Section 4, we shall present \textit{a posteriori} error estimator and prove the discrete upper bound. In Section 5, we shall present our algorithms and prove their convergence and optimality.

Throughout this paper, we shall use standard notation for Sobolev spaces and use boldface letter for the spaces of vectors. The letter \(C\), without subscript, denotes generic constants that may not be the same at different occurrences and \(C_i\), with subscript, denotes specific important constants.

2. Preliminaries

In this section we shall introduce mixed finite element methods for the Poisson equation and discuss the general procedure of adaptive methods through local refinement. We shall also include a result on the approximation of the data.
2.1. **Mixed finite element methods.** The standard finite element method involves writing (1.1) as a primary variational formulation: for a given \( f \in L^2(\Omega) \), find \( u \in H^1_0(\Omega) \) such that

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

and then finding an approximation by solving (2.1) in finite-dimensional subspaces of \( H^1_0(\Omega) \). In many applications ([23]) the variable \( \sigma = -\nabla u \) is of interest, and it is therefore convenient to use mixed finite element methods. Let us first write (1.1) as a first order system:

\[
\sigma + \nabla u = 0, \quad \text{div } \sigma = f \quad \text{in } \Omega, \quad \text{and } u = 0 \quad \text{on } \partial \Omega.
\]

Let

\[
\Sigma = H(\text{div}; \Omega) := \{ \tau \in L^2(\Omega) : \text{div } \tau \in L^2(\Omega) \}, \quad \text{and } U = L^2(\Omega).
\]

We shall use \( \| \cdot \| \) to denote \( L^2 \)-norm and \( \| \cdot \|_{H(\text{div})} \) for the \( H(\text{div}) \) norm:

\[
\| \tau \|_{H(\text{div})} = (\| \tau \|^2 + \| \text{div } \tau \|^2)^{1/2}, \quad \forall \tau \in \Sigma.
\]

The mixed (or dual) variational formulation of (2.2) is, given an \( f \in L^2(\Omega) \), find \((\sigma, u) \in \Sigma \times U\) such that

\[
(\sigma, \tau) - (\text{div } \tau, u) = 0 \quad \forall \tau \in \Sigma,
\]

\[
(\text{div } \sigma, v) = (f, v) \quad \forall v \in U,
\]

where \((\cdot, \cdot)\) is the inner product for \( L^2(\Omega) \) or \( L^2(\Omega) \). Note that the Dirichlet boundary condition is imposed as a natural boundary condition in the dual formulation (2.3) using integration by parts. The existence and uniqueness of the solution \((\sigma, u)\) to (2.3)-(2.4) follows from the so-called inf-sup condition which can be easily established for this model problem [23].

Given a shape regular and conforming (in the sense of [28]) triangulation \( T_H \) of \( \Omega \), the mixed finite element method is to solve (2.3)-(2.4) in a pair of finite-dimensional spaces \( \Sigma_H \subset \Sigma \) and \( U_H \subset U \). That is, given an \( f \in L^2(\Omega) \), to find \((\sigma_H, u_H) \in \Sigma_H \times U_H\) such that

\[
(\sigma_H, \tau_H) - (\text{div } \tau_H, u_H) = 0 \quad \forall \tau_H \in \Sigma_H
\]

\[
(\text{div } \sigma_H, v_H) = (f_H, v_H) \quad \forall v_H \in U_H.
\]

Hereafter \( f_H \) denotes the \( L^2(\Omega) \) projection of \( f \) onto \( U_H \). Namely, \( f_H \in U_H \) such that \((f_H, v_H) = (f, v_H), \forall v_H \in U_H\). The well-posedness of the discrete problem (2.5)-(2.6), unlike the standard finite element method for the primary variational formulation, is non-trivial. One sufficient condition to construct stable finite element spaces is to ensure the inf-sup condition still holds for the discrete problem. Since 1970’s many stable finite element spaces have been introduced for this case, such as those of Raviart-Thomas spaces [49] and Brezzi-Douglas-Marini spaces [22]. Recently it has been shown that such stable finite element spaces can be constructed in an elegant way using differential complex theory [15, 39, 2, 4].

The Raviart-Thomas spaces [49] are defined for an integer \( p \geq 0 \) by

\[
RT_H = \Sigma_H^p \times U_H^p,
\]

where

\[
\Sigma_H^p(T) := \{ \tau \in H(\text{div}; T) : \tau|_T \in P_p(T) + xP_p(T), \forall T \in T_H \},
\]

and

\[
U_H^p(T) := \{ v \in L^2(\Omega) : v|_T \in P_p(T), \forall T \in T_H \}.
\]
and where $P_p(T)$ denotes the space of polynomials on $T$ of degree at most $p$.

The Brezzi-Douglas-Marini spaces [22] are defined for an integer $p \geq 1$ by

$$BDM_H = \Sigma_H^p \times U_H^p,$$

where

$$\Sigma_H^p(T_H) := \{ \tau \in H(\text{div}; \Omega) : \tau|_T \in P_p(T), \forall T \in T_H \},$$

and $U_H^p(T_H) := \{ v \in L^2(\Omega) : v|_T \in P_{p-1}(T), \forall T \in T_H \}$.

Since most results hold for both Raviart-Thomas and Brezzi-Douglas-Marini spaces and $p$ is fixed in most places, we shall use generic notation $(\Sigma_H, U_H)$ to denote the pair in $RT_H$ or $BDM_H$. The discrete problem posed on $(\Sigma_H, U_H)$ will satisfy the discrete inf-sup condition [23] from which the existence and uniqueness of the finite element approximation $(\sigma_H, u_H)$ follows.

We shall use $\mathcal{L}$ and $\mathcal{L}_H$ to denote the differential operators corresponding to (2.3)-(2.4) and (2.5)-(2.6), respectively. Those equations can be formally written as

$$\mathcal{L}(\sigma, u) = f \quad \text{and} \quad \mathcal{L}_H(\sigma_H, u_H) = f_H.$$

We shall use the notation $(\sigma, u) = \mathcal{L}^{-1} f$ and $(\sigma_H, u_H) = \mathcal{L}_H^{-1} f$ to emphasis the dependence of $f$. With an abuse of notation, we also use $\sigma = \mathcal{L}^{-1} f$ and $\sigma_H = \mathcal{L}_H^{-1} f_H$ when $\sigma$ and $\sigma_H$ are of interest.

### 2.2. Adaptive methods through local refinement.

Let $\sigma = \mathcal{L}^{-1} f$ and $\sigma_H = \mathcal{L}_H^{-1} f_H$.

We are mostly interested in the control of the error $\| \sigma - \sigma_H \|$ which is usually more important than control the error of scalar variable $u$ in mixed finite element methods. Although the natural norm for the error is $\| \sigma - \sigma_H \|_{H(\text{div})}$, we comment that, by (2.4) and (2.6), $\| \text{div } \sigma - \text{div } \sigma_H \| = \| f - f_H \|$ can be approximated efficiently without solving equations and also may dominate the error $\| \sigma - \sigma_H \|_{H(\text{div})}$; see Remark 3.4 in [42].

The rate of the error $\| \sigma - \sigma_H \|$ for $\sigma_H \in \Sigma_H^p(T_H)$ depends on the regularity of the function being approximated and the regularity of the mesh. If $\sigma \in H^{p+1}(\Omega)$ and $T_H$ is quasi-uniform with mesh size $H = \max_{T \in T_H} \text{diam}(T)$, then the following convergence result of optimal order is well known [23]

$$\| \sigma - \sigma_H \| \leq CH^{p+1} \| \sigma \|_{p+1}. \tag{2.7}$$

The regularity result $\sigma \in H^{p+1}(\Omega)$, however, may not be true in many applications, especially for concave domains $\Omega$. Thus we cannot expect the convergence result (2.7) on quasi-uniform grids in general.

To improve the convergence rate, element sizes are adapted according to the behavior of the solution. In this case, the element size in areas of the domain where the solution is smooth can stay bounded well away from zero, and thus the global element size is not a good measure of the approximation rate. For this reason, when the optimality of the convergence rate is concerned, $\# T$, the number of elements, is used to measure the approximation rate in the setting of adaptive methods that involve local refinement.

We now briefly review the standard adaptive procedure. Given an initial triangulation $T_0$, we shall generate a sequence of nested conforming triangulation $T_k$ using the following loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \tag{2.8}$$

More precisely, to get $T_{k+1}$ from $T_k$ we first solve (2.5)-(2.6) to get $\sigma_k$ on $T_k$. The error is estimated using $\sigma_k$ and data. And the error estimator is used to mark a set of of triangles or edges that are to be refined. Triangles are then refined in such a way that the triangulation is still shape regular and conforming in the sense of [28].
We shall not discuss the step **SOLVE** which deserves a separate investigation. We assume that the solutions of the finite-dimensional problems can be generated to any accuracy to accomplish this in optimal space and time complexity. Multigrid-like methods for mixed finite element methods on quasi-uniform grids can be found in [16, 17, 19, 20, 38, 48, 52].

The *a posteriori* error estimators are essential part of the **ESTIMATE** step. Given a shape regular triangulation $T_H$, let $E_H$ denotes the edges of $T_H$. In this paper, we shall use edge-wised error estimator $\eta_E$ for each edge $E \in E_H$. See Section 4 for details.

The local error estimator $\eta_E$ is employed to mark for refinement the elements whose error estimator is large. The way we mark these triangles influences the efficiency of the adaptive algorithm. In the **MARK** step we shall always use the marking strategy firstly proposed by Dörfler [34] in order to prove the convergence and the optimality of the local refinement strategy.

In the **REFINE** step we need to carefully choose the rule for dividing the marked triangles such that the mesh obtained by this dividing rule is still conforming and shape regular. Such refinement rules include red and green refinement [10], longest edge refinement [51, 50], and newest vertex bisection [54]. Note that not only marked triangles get refined but also additional triangles are refined to recovery the conformity of triangulations. We would like to control the number of elements added to ensure the overall optimality of the refinement procedure. To this end, we shall use newest vertex bisection in this article. We refer to [43, 57, 11, 27] for details of newest vertex bisection and only list two important properties below.

Let $T_k$ be a conforming triangulation refined from a shape regular triangulation $T_0$ using the newest vertex bisection and let $\mathcal{M}$ be the collection of all marked triangles going from $T_0$ to $T_k$. Then

1. $\{T_k\}$ is uniform shape regular with respect to $k$ and the shape regularity only depends on $T_0$;
2. $\# T_k \leq \# T_0 + C \# M$.

Recently Stevenson [56] shows that such results can be extended to bisection algorithms of $n$-simplices. The optimality of the adaptive finite element method in this paper, thus, could be easily extended to general space dimensions.

### 2.3. Approximation of the data

We shall introduce the concept of data oscillation which is firstly introduced in [44], and use it here for the approximation of data. Such quantity measures intrinsic information missing in the averaging process associated with finite elements, which fails to detect fine structures of $f$.

For a set $A$, $H_A$ denotes the diameter of $A$. To easy the notation, we may drop the subscript if it is clear from the context. For a triangulation $T_H$ of $\Omega$ and a function $f \in L^2(\Omega)$, we define a triangulation dependent norm

$$
\| H f \|_{0,T_H} := \left( \sum_{T \in T_H} H_T \| f \|_{0,T}^2 \right)^{1/2}.
$$

**Definition 2.1.** Given a shape regular triangulation $T_H$ of $\Omega$ and an $f \in L^2(\Omega)$, we define the data oscillation

$$
\text{osc}(f, T_H) := \| H (f - f_{T_H}) \|_{0,T_H}.
$$

Let $\mathcal{P}_N$ denote the set of triangulations constructed from an initial triangulation $T_0$ by the newest vertex bisection method with at most $N$ triangles. We define

$$
\| f \|_{\mathcal{A}_N} = \sup_{N \geq N_0} \left( N^s \inf_{T \in \mathcal{P}_N} \text{osc}(f, T) \right),
$$

where $s \in (0, 1)$.
where \( N_0 \) is a fixed integer representing the number of triangles in \( T_0 \). We will recall a result of Binev, Dahmen and DeVore [11] which shows that the approximation of data can be done in an optimal way. The proof can be found at [11]; See also [13].

**Theorem 2.2** (Binev, Dahmen and DeVore). Given a tolerance \( \varepsilon \), an \( f \in L^2(\Omega) \) and a shape regular triangulation \( T_0 \), there exists an algorithm

\[
T_H = \text{APPROX}(f, T_0, \varepsilon)
\]

such that

\[
\text{osc}(f, T_H) \leq \varepsilon, \quad \text{and} \quad \#T_H - \#T_0 \leq C\|f\|^{1/s} h_0^{1/s} \varepsilon^{-1/s}.
\]

### 3. Quasi-Orthogonality

Unlike the primary weak form of Poisson equation, \( \sigma_H \) is not the \( L^2 \)-orthogonal projection of \( \sigma \) from \( \Sigma \) to \( \Sigma_H \). Indeed the solution \((\sigma, u)\) of (2.3)-(2.4) is the saddle point of the following energy

\[
E(\tau, v) = \frac{1}{2} \|\tau\|^2 + (\text{div} \, \tau, v) - (f, v), \quad \tau \in H(\text{div}; \Omega), \quad v \in L^2(\Omega).
\]

Namely

\[
E(\sigma, u) = \inf_{\sigma \in H(\text{div}; \Omega)} \sup_{v \in L^2(\Omega)} E(\tau, v).
\]

Similar result holds for the discrete solutions \((\sigma_H, u_H)\). The lack of orthogonality is the main difficulty which complicates the convergence analysis of mixed finite element methods.

We shall use the fact the error \( \sigma - \sigma_H \) is orthogonal to the divergence free subspace of \( \Sigma_H \) to prove a quasi-orthogonality result. In the sequel we shall consider two conforming triangulations \( T_h \) and \( T_H \) which are nested in the sense that \( T_h \) is a refinement of \( T_H \). Therefore the finite element space are nested i.e. \((\Sigma_H, U_H) \subset (\Sigma_h, U_h)\).

**Lemma 3.1.** Given an \( f \in L^2(\Omega) \) and two nested triangulation \( T_h \) and \( T_H \), let

\[
(\sigma, u) = L^{-1} f, \quad (\sigma_h, u_h) = L_h^{-1} f_h, \quad (\tilde{\sigma}, \tilde{u}) = L^{-1} f_H, \quad \text{and} \quad (\sigma_H, u_H) = L_H^{-1} f_H.
\]

Then

\[
(\sigma - \sigma_h, \sigma_h - \sigma_H) = 0.
\]

**Proof.** Since \( \tilde{\sigma} - \tilde{\sigma}_h \in \Sigma_h \), by (2.5)-(2.6), we have

\[
(\sigma - \sigma_h, \tilde{\sigma} - \tilde{\sigma}_H) = (u - u_h, \text{div} (\tilde{\sigma} - \tilde{\sigma}_H)) = (u - u_h, f_H - f_H) = 0.
\]

\( \square \)

To prove quasi-orthogonality, we need the following discrete stability result

\[
(\sigma - \sigma_h, \sigma_h - \sigma_H) \leq \sqrt{C_0} \, \text{osc}(f_h, T_H),
\]

where the constant \( C_0 \) depends only on \( T_H \). We shall leave the proof of (3.2) to the next section and use it to derive the quasi-orthogonality result.

**Theorem 3.2.** Given an \( f \in L^2(\Omega) \) and two nested triangulation \( T_h \) and \( T_H \), let \( \sigma = L^{-1} f, \sigma_h = L_h^{-1} f_h \), and \( \sigma_H = L_H^{-1} f_H \). Then

\[
(\sigma - \sigma_h, \sigma_h - \sigma_H) \leq \sqrt{C_0} \|\sigma - \sigma_h\| \, \text{osc}(f_h, T_H),
\]
Thus for any $\delta > 0$, 

\begin{equation}
(1-\delta)\|\sigma - \sigma_h\|^2 \leq \|\sigma - \sigma_H\|^2 - \|\sigma_h - \sigma_H\|^2 + \frac{C_0}{\delta}\operatorname{osc}^2(f_h, T_H),
\end{equation}

and in particular when $\operatorname{osc}(f_h, T_H) = 0$,

\begin{equation}
\|\sigma - \sigma_h\|^2 = \|\sigma - \sigma_H\|^2 - \|\sigma_h - \sigma_H\|^2.
\end{equation}

**Proof.** Let us introduce an intermediate solution $\tilde{\sigma}_h = L^{-1}_h f_h$. By Lemma 3.1, $(\sigma - \sigma_h, \tilde{\sigma}_h - \sigma_H) = 0$. Thus 

$(\sigma - \sigma_h, \sigma_h - \sigma) = (\sigma - \sigma_h, \sigma_h - \tilde{\sigma}_h) \leq \|\sigma - \sigma_h\|\|\sigma_h - \tilde{\sigma}_h\|$.

(3.3) then follows from the inequality (3.2).

By the trivial identity $\sigma - \sigma_H = \sigma - \sigma_h + \sigma_h - \sigma_H$, we have

$$
\|\sigma - \sigma_H\|^2 = \|\sigma - \sigma_h\|^2 + 2(\sigma - \sigma_h, \sigma_h - \sigma_H)
$$

When $\operatorname{osc}(f_h, T_H) = 0$, by (3.3), $(\sigma - \sigma_h, \sigma_h - \sigma_H) = 0$ and thus (3.5) follows. In general, we use

$$
\|\sigma - \sigma_H\|^2 = \|\sigma - \sigma_h\|^2 + \|\sigma_h - \sigma_H\|^2 + 2(\sigma - \sigma_h, \sigma_h - \sigma_H)
$$

\begin{align*}
&\geq \|\sigma - \sigma_h\|^2 + \|\sigma_h - \sigma_H\|^2 - 2\sqrt{C_0}\|\sigma - \sigma_h\|\operatorname{osc}(f, T_H) \\
&\geq \|\sigma_h - \sigma_H\|^2 + (1-\delta)\|\sigma - \sigma_h\|^2 - \frac{C_0}{\delta}\operatorname{osc}^2(f_h, T_H),
\end{align*}

to prove (3.4). In the last step, we have used the inequality

$$
2ab \leq \delta a^2 + \frac{1}{\delta} b^2.
$$

\[\square\]

A similar quasi-orthogonality result was obtained by Carstensen and Hoppe [25] for the lowest order Raviart-Thomas spaces using a special relation to the non-conforming finite element. Such relation for high order elements and Brezzi-Douglas-Marini spaces are not easy to establish; see [3] and [29, 30, 31] for discussion on this relation. In contrast the approach we used here is more straight-forward.

**Remark 3.3.** The oscillation term $\operatorname{osc}(f_h, T_H)$ in (3.3) and (3.4) depends on both $T_h$ and $T_H$. It can be changed to the quantity $\operatorname{osc}(f, T_H)$ which only depends on $T_H$. Indeed for each $T \in T_H$, we have

$$
\|f_h - f_H\|_{0,T} = \|Q_h(I - Q_H) f\|_{0,T} \leq \|f - f_H\|_{0,T},
$$

and thus $\operatorname{osc}(f_h, T_H) \leq \operatorname{osc}(f, T_H)$. This change is important for the construction of convergent AMFEM by showing the reduction of $\operatorname{osc}(f, T_H)$.

4. **Discrete stability for perturbation of data**

In this section, we shall prove the discrete stability result. Let us first prove a stability result in the continuous case. Let $u \in H^1_0(\Omega)$ be the solution of the primary weak form (2.1) of Poisson equation. Then $(-\nabla u, u)$ is the solution to the dual weak form (2.3)-(2.4). The following stability result $\|\sigma\| \leq \|f\|_{-1}$ is well-known in the literature. The norm $\|f\|_{-1}$, however, is not easy to compute. Instead we shall make use of the oscillation of data to bound it.
Theorem 4.1. Given a shape regular triangulation $T_H$ of $\Omega$ and $f \in L^2(\Omega)$, let $(\sigma, u) = \mathcal{L}^{-1} f$ and $(\tilde{\sigma}, \tilde{u}) = \mathcal{L}^{-1} f_H$, respectively. Then there exists a constant $C_0$ depending only on the shape regularity of $T_H$ such that

\begin{equation}
\| \sigma - \tilde{\sigma} \| \leq \sqrt{C_0 \operatorname{osc}(f, T_H)}.
\end{equation}

Proof. By (2.3) and (2.4), we have

\[ \| \sigma - \tilde{\sigma} \|^2 = (\sigma - \tilde{\sigma}, \sigma - \tilde{\sigma}) = (\text{div}(\sigma - \tilde{\sigma}), u - \tilde{u}) = (f - f_H, u - \tilde{u}). \]

Let $v$ be the solution of primary weak formulation of Poisson equation with data $f - f_H$. Then $v = u - \tilde{u}$ and $-\nabla v = \sigma - \tilde{\sigma}$. Let $Q_H : L^2(\Omega) \to U_H$ be the $L^2$ projection. Since $f_H = Q_H f$, we have

\[ \| \sigma - \tilde{\sigma} \|^2 = (f - f_H, v) \]

\[ = \sum_{T \in T_H} (f - f_H, v - Q_H v)_T \]

\[ \leq \sqrt{C_0} \sum_{T \in T_H} \| H(f - f_H) \|_{0,T} \| \nabla v \|_{0,T} \]

\[ \leq \sqrt{C_0} \left( \sum_{T \in T_H} \| H(f - f_H) \|^2_{0,T} \right)^{1/2} \| \sigma - \tilde{\sigma} \|. \]

In the second step, we have used the error estimate

\[ \| v - Q_H v \|_{0,T} \leq \sqrt{C_0 H_T} \| \nabla v \|_{0,T}, \]

which can be easily proved by Bramble-Hilbert lemma and the scaling argument. The constant $C_0$ only depends on the shape regularity of $T_H$. The desired result then follows by canceling one $\| \sigma - \tilde{\sigma} \|$. \hfill \square

In the proof of Theorem 4.1, we use the local error estimate

\[ \| u - Q_H u \|_{0,T} \leq \sqrt{C_0 H_T} \| \nabla u \|_{0,T} = \sqrt{C_0 H_T} \| \sigma \|_T, \]

for $u \in H^1_0(\Omega)$ and $\sigma = -\nabla u$. The main difficulty in the discrete case is that $u_h \in U_h \notin H^1_0(\Omega)$. However we still have a similar localized error estimate for $u_h - Q_H u_h$.

Lemma 4.2. Let $T_h$ and $T_H$ be two nested triangulations, and let $(\sigma_h, u_h) = \mathcal{L}^{-1} f_h$. Then for any $T \in T_H$, we have

\begin{equation}
\| u_h - Q_H u_h \|_{0,T} \leq \sqrt{C_0 H_T} \| \sigma_h \|_{0,T}. \end{equation}

The proof of this lemma is technical and postponed to the end of this section. We use it to prove the following theorem.

Theorem 4.3. Let $T_h$ and $T_H$ be two nested conforming triangulations. Let $\tilde{\sigma}_h = \mathcal{L}_h^{-1} f_H$ and $\sigma_h \in \mathcal{L}_h^{-1} f_h$. Then there exists a constant $C_0$, depending only on the shape regularity of $T_H$ such that

\begin{equation}
\| \sigma_h - \tilde{\sigma}_h \| \leq \sqrt{C_0 \operatorname{osc}(f_h, T_H)}.
\end{equation}

Proof. Recall that $\sigma_h - \tilde{\sigma}_h$ satisfies the equation

\begin{equation}
(\sigma_h - \tilde{\sigma}_h, \tau_h) = (u_h - \tilde{u}_h, \text{div} \tau_h), \quad \forall \tau_h \in \Sigma_h
\end{equation}

\begin{equation}
(\text{div}(\sigma_h - \tilde{\sigma}_h), v_h) = (f_h - f_H, v_h), \quad \forall v_h \in U_h.
\end{equation}

We then choose $\tau_h = \sigma_h - \tilde{\sigma}_h$ in (4.4) and $v_h = u_h - \tilde{u}_h$ in (4.5) to obtain

\[ \| \sigma_h - \tilde{\sigma}_h \|^2 = (u_h - \tilde{u}_h, \text{div}(\sigma_h - \tilde{\sigma}_h)) = (v_h, f_h - f_H) = (v_h - Q_H v_h, f_h - f_H). \]
In the third step, we use the fact $f_H = Q_H f = Q_H f_h$ since $T_h$ and $T_H$ are nested. Then using (4.2), we have
\[
\|\sigma_h - \tilde{\sigma}_h\|^2 = \sum_{T \in T_H} (v_h - Q_H v_h, f_h - f_H)_T \\
\leq \sqrt{C_0} \sum_{T \in T_H} H_T \|f_h - f_H\|_{0,T} \|\sigma_h - \tilde{\sigma}_h\|_{0,T} \\
\leq \sqrt{C_0} \left( \sum_{T \in T_H} H_T^2 \|f_h - f_H\|_T^2 \right)^{1/2} \|\sigma_h - \tilde{\sigma}_h\|.
\]
Canceling one $\|\sigma_h - \tilde{\sigma}_h\|$, we get the desired result. \hfill \Box

In the rest of this section, we shall prove Lemma 4.2. The first ingredient is the existence of a continuous right inverse of the divergence as an operator from $H^1_0(\Omega)$ into the space $L^2_0(\Omega) := \{ v \in L^2(\Omega) : \int_\Omega v = 0 \}.$

**Lemma 4.4.** Given a function $f \in L^2_0(\Omega)$, there exists a function $\tau \in H^1_0(\Omega)$ such that
\[
\text{div} \, \tau = f \quad \text{and} \quad \|\tau\|_1 \leq C\|f\|.
\]

The proof of this lemma for smooth or convex domains $\Omega$ is pretty easy. One can solve the Poisson equation with Neumann boundary condition
\[
\Delta \phi = f, \quad \text{in } \Omega, \quad \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \partial \Omega.
\]
The condition $f \in L^2_0(\Omega)$ ensures the existence of the solution. Then we let $\tau = \text{grad} \, \phi$ and modify the tangent component of $\tau$ to be zero [21]. See also [6, 35] for detailed proof on non-convex and general Lipschitz domains.

The second ingredient is an interpolation operator $\Pi_h : H^1(\Omega) \rightarrow \Sigma_h$ with the following nice properties.

**Lemma 4.5.** There exists an interpolation operator $\Pi_h : H^1(\Omega) \rightarrow \Sigma_h$ such that
\begin{enumerate}
  
  \item $Q_h \text{div} \, \tau = \text{div} \, \Pi_h \tau, \quad \forall \tau \in H^1(\Omega)$;
  
  \item there exists a constant $C$ only depending on the shape regularity of $T_h$ such that
    \[
    \|\tau - \Pi_h \tau\|_T \leq C h_T \|\tau\|_{1,T}, \quad \forall T \in T_h, \forall \tau \in H^1(\Omega);
    \]
  
  \item for any $T \in T_h$ if $\tau \in H^1_0(T)$, then $\Pi_h \tau|_{\partial T} = 0$.
\end{enumerate}

For the detailed construction of such interpolation operator and proof of these properties, we refer to [39] and [5].

**Proof of Lemma 4.2** We first note that $u_h - Q_H u_h = (Q_h - Q_H)u_h$ since $Q_h u_h = u_h.$ For any $T \in T_H$, by the definition of $L^2$ projection $Q_H$, we have, $\int_T (Q_h - Q_H)u_h = 0$ i.e. $(Q_h - Q_H)u_h \in L^2_0(T).$ We thus can apply Lemma 4.4 to find a function $\tau \in H^1_0(T)$ such that
\[
\text{div} \, \tau = (Q_h - Q_H)u_h, \quad \text{in } T \quad \text{and} \quad \|\tau\|_{1,T} \leq C \|(Q_h - Q_H)u_h\|_{0,T}.
\]

We extend $\tau$ to $H^1(\Omega)$ by zero. Note that
\[
(\Pi_h - Q_H)\tau \in \Sigma_h, \quad \text{and sup}\,(\Pi_h - Q_H)\tau \subseteq T.
\]

With such $\tau$, we have
\[
\|(Q_h - Q_H)u_h\|_{0,T}^2 = (Q_h - Q_H)u_h, \text{div} \, \tau)_T = (u_h, (Q_h - Q_H) \text{div} \, \tau)_T.
\]
Then using the commuting property (Lemma 4.5 (1)) and the locality of $\tau$, we have
\[(u_h, (Q_h - Q_H) \text{div} \tau)_T = (u_h, (Q_h - Q_H) \text{div} \tau)_\Omega = (u_h, \text{div}(\Pi_h - \Pi_H)\tau)_\Omega.\]
Now we shall use the fact $(\sigma_h, u_h)$ is the solution of (2.3) and (2.4) and, again, the locality of $\tau$ to get
\[(u_h, \text{div}(\Pi_h - \Pi_H)\tau)_\Omega = (\sigma_h, (\Pi_h - \Pi_H)\tau)_\Omega = (\sigma_h, (\Pi_h - \Pi_H)\tau)_T.\]
Using the approximation property of $\Pi_h$, we get
\[\|\sigma_h, (\Pi_h - \Pi_H)\tau\|_T \leq C_H \|\sigma_h\|_0, T \|\tau\|_{1, T}.\]
So we have
\[\|Q_h - Q_H\|_{0, T} \leq C_H \|\sigma_h\|_0, T \|\tau\|_{1, T}.\]
Canceling one term $(Q_h - Q_H)u_h\|_0, T$, we get the desired result. \hfill \square

5. A Posteriori Error Estimate for Mixed Finite Element Methods

In this section we shall follow Alonso [1] to present a posteriori error estimate for mixed finite element methods. Other a posteriori error estimators for the mixed finite element methods can be found at [24, 58, 40, 37, 41, 42]. Our analysis could be adapted to these error estimators also.

5.1. A posteriori error estimator and existing results. Let us begin with the definition of the error estimator. For any edge $E \in \mathcal{E}_h$, we shall fix an unit tangent vector $t_E$ for $E$. We denote the patch of $E$ consisting of triangles sharing $E$ by $\Omega_E$.

**Definition 5.1.** Given a triangulation $T_h$, for an $E \in \mathcal{E}_h$ and $E \notin \partial \Omega$, let $\Omega_E = T \cup \bar{\bar{T}}$. For any $\sigma_h \in \Sigma_h$, we define the jump of $\sigma_h$ across edge $E$ as
\[J_E(\sigma_h) = [\sigma_h \cdot t_E] := \sigma_h|_T \cdot t_E - \sigma_h|_{\bar{T}} \cdot t_E.\]
If $E \in \mathcal{E}_h \cap \partial \Omega$, we define $J_E(\sigma_h) = \sigma_h \cdot t_E$. The edge error estimator is defined as
\[\eta^2_E(\sigma_h) = \|H \text{rot} \sigma_h\|_{0, \Omega_E}^2 + \|H^{1/2} J_E(\sigma_h)\|_{0, E}^2.\]
For a subset $\mathcal{F}_h \subseteq \mathcal{E}_h$, we define
\[\eta^2(\sigma_h, \mathcal{F}_h) := \sum_{E \in \mathcal{F}_h} \eta^2_E(\sigma_h).\]

The error estimator $\eta_E(\sigma_h)$ is continuous with respect to $\sigma_h$ in $L^2$-norm. Namely we have the following inequality.

**Lemma 5.2.** Given an $f \in L^2(\Omega)$ and a shape regular triangulation $T_h$, let $\sigma_h, \tau_h \in \Sigma_h$. There exists constant $\beta$ such that
\[\beta \|\eta^2(\sigma_h, \mathcal{E}_h) - \eta^2(\tau_h, \mathcal{E}_h)\| \leq \|\sigma_h - \tau_h\|^2.\]
Proof. It can be easily proved by the triangle inequality and inverse inequality. \hfill \square

We shall recall Alonso’s results below and prove a discrete upper bond later. Since the data $f$ is not included in the definition of our error estimator $\eta_E$, the upper bound contains additional data oscillation term which is different with standard one in [57].
Theorem 5.3 (Upper bound). Given an $f \in L^2(\Omega)$ and a shape regular triangulation $T_H$, let $\sigma = L^{-1} f$ and $\sigma_H = L^{-1}_H f_H$. There exist constants $C_0$ and $C_1$ depending only the shape regularity of $T_H$ such that
\begin{equation}
\|\sigma - \sigma_H\|^2 \leq C_1 \eta^2(\sigma_H, E_H) + C_0 \text{osc}^2(f, T_H).
\end{equation}

Theorem 5.4 (Lower bound). Given an $f \in L^2(\Omega)$ and a shape regular triangulation $T_H$, let $\sigma = L^{-1} f$ and $\sigma_H = L^{-1}_H f_H$. There exists constant $C_2$ depending only the shape regularity of $T_H$ such that
\begin{equation}
C_2 \eta^2(\sigma_H, E_H) \leq \|\sigma - \sigma_H\|^2,
\end{equation}
for Raviart-Thomas spaces. For Brezzi-Douglas-Marini spaces, (5.4) holds when $\text{osc}(f, T_H) = 0$.

When $\text{osc}(f, T_H) = 0$, (5.3) and (5.4) implies that $C_2/C_1 \leq 1$. This ratio is a measure of the precision of the indicator.

5.2. Discrete upper bound. We shall give a discrete version of the upper bound (5.3). The main tool is the discrete Helmholtz decomposition.

Given a shape regular triangulation $T_h$, let
\[ S_h^p = \{ \psi_h \in C(\Omega) : \psi_h|_T \in P_p(T), \forall T \in T_h \} \]
denote the standard continuous and piecewise polynomial finite element spaces of $H^1(\Omega)$.

To introduce the discrete Helmholtz decomposition, we define the dual operator operator $\text{div} : \Sigma_h \mapsto U_h$.

Definition 5.5. We define $\text{grad}_h : U_h \mapsto (\Sigma_h)^*$ by
\[ (\text{grad}_h v_h, \tau_h) = (v_h, \text{div} \tau_h), \quad \forall \tau_h \in \Sigma_h. \]

We emphasis that $\text{grad}_h$ is not simply the restriction of $\text{grad}$ to $U_h$ since the integration by parts will give non-vanishing boundary term. The following discrete Helmholtz decomposition is well known in the literature; See, for example, [36, 18, 4, 14].

Theorem 5.6 (Discrete Helmholtz Decomposition in $\mathbb{R}^2$). Given a triangulation $T_h$, for $p$-th order Raviart-Thomas finite element spaces $(\Sigma_h^p, U_h^p)$, we have the following orthogonal (with respect to $L^2$ inner product) decomposition
\[ \Sigma_h^p = \text{curl}(S_h^{p+1}) \oplus \text{grad}_h(U_h^p). \]

For Brezzi-Douglas-Marini finite element spaces $(\Sigma_h^p, U_h^p)$, we have the following orthogonal (with respect to $L^2$ inner product) decomposition
\[ \Sigma_h^p = \text{curl}(S_h^{p+1}) \oplus \text{grad}_h(U_h^{p-1}). \]

We are in the position to present a discrete version of the upper bound.

Theorem 5.7. Let $T_h$ and $T_H$ be two nested conforming triangulations. Let $\sigma_h = L^{-1}_h f_h$ and $\sigma_H = L^{-1}_H f_H$, and let $\mathcal{F}_H = \{ E \in E_H : E \notin E_h \}$. Then there exist constants depending only the shape regularity of $T_H$ such that
\begin{equation}
\|\sigma_h - \sigma_H\|^2 \leq C_1 \eta^2(\sigma_H, E_H) + C_0 \text{osc}^2(f_h, T_H)
\end{equation}
and
\begin{equation}
\# \mathcal{F}_H \leq 3(\# T_h - \# T_H).
\end{equation}
Proof. The inequality (5.6) follows from

$$\# \mathcal{F}_H \leq \# \mathcal{E}_h - \# \mathcal{E}_H \leq 3(\# T_h - \# T_H).$$

To prove (5.5), again we introduce the intermediate solution \( \tilde{\sigma}_h = L_h^{-1} f_H \). By the discrete Helmholtz decomposition, we have

$$\tilde{\sigma}_h - \sigma_H = \text{grad}_h \phi_h + \text{curl} \psi_h,$$

where \( \phi_h \in U_h^p, \psi_h \in S_h^{p+1} \) for Raviart-Thomas spaces, and \( \phi_h \in U_h^{p-1}, \psi_h \in S_h^{p+1} \), for Brezzi-Douglas-Marini spaces. The decomposition is \( L^2 \)-orthogonal i.e.

$$\| \tilde{\sigma}_h - \sigma_H \|^2 = \| \text{grad}_h \phi_h \|^2 + \| \text{curl} \psi_h \|^2.$$  \hfill (5.7)

In two dimensions, \( \| \text{curl} \psi_h \| = \| \text{grad} \psi_h \| \) and thus (5.7) implies that

$$\| \psi_h \| \leq \| \tilde{\sigma}_h - \sigma_H \|.$$  \hfill (5.8)

Since

$$\langle \sigma_h - \sigma_H, \text{grad}_h v_h \rangle = \langle \text{div} (\sigma_h - \sigma_H), v_h \rangle = \langle f_H - f_H, v_h - v_H \rangle = 0,$$

we have

$$\| \tilde{\sigma}_h - \sigma_H \|^2 = \langle \tilde{\sigma}_h - \sigma_H, \text{grad}_h \phi_h \rangle + \langle \tilde{\sigma}_h - \sigma_H, \text{curl} \psi_h \rangle = \langle \tilde{\sigma}_h - \sigma_H, \text{curl} \psi_h \rangle.$$  \hfill (5.9)

Now we write \( \sigma_h - \sigma_H = \sigma_h - \tilde{\sigma}_h + \tilde{\sigma}_h - \sigma_H \) and note that

$$\langle \sigma_h - \tilde{\sigma}_h, \tilde{\sigma}_h - \sigma_H \rangle = \langle u_h - \tilde{u}_h, \text{div}(\tilde{\sigma}_h - \sigma_H) \rangle = \langle u_h - \tilde{u}_h, f_H - f_H \rangle = 0.$$

Combining (5.9) and (3.2), we then have

$$\| \sigma_h - \sigma_H \|^2 = \| \tilde{\sigma}_h - \sigma_H \|^2 + \| \sigma_h - \tilde{\sigma}_h \|^2 \leq C_1 \eta(\sigma_H, \mathcal{F}_H) + \mathcal{C}_0 \text{osc}^2(f_h, T_H).$$

□
6. CONVERGENCE AND OPTIMALITY OF AMFEM

In this section we shall present our algorithms and prove their convergence and optimality. It is adapted from the literature [44, 55, 34, 45, 46]. For the completeness we include them here and prove some important technique results.

We first present our algorithms. It mainly follows from the algorithm proposed in [46]. The difference is that we do not impose interior point property, in the refinement.

Let $T_0$ be a initial shape regular triangulation, a right side $f \in L^2(\Omega)$, a tolerance $\varepsilon$, and $0 < \theta, \tilde{\theta}, \mu < 1$ three parameters. Thereafter we replace the subscript $h$ by an iteration counter called $k$. For a marked edge set $M_k$, we denote by $\Omega_{M_k} = \bigcup_{E \in M_k} \Omega_E$.

$[T_N, \sigma_N] = AMFEM(T_0, f, \varepsilon, \theta, \tilde{\theta}, \mu)$

$\eta = \varepsilon, k = 0$

WHILE $\eta \geq \varepsilon$, DO

Solve (2.5)-(2.6) on $T_k$ to get the solution $\sigma_k$.

Compute the error estimator $\eta = \eta(\sigma_k, E_k)$.

Mark the minimal edge set $M_k$ such that

(6.1) $\eta^2(\sigma_k, M_k) \geq \theta \eta^2(\sigma_k, E_k)$.  

If $\text{osc}(f, T_k) > \text{osc}(f, T_0) \mu^k$, enlarge $M_k$ such that

(6.2) $\text{osc}(f, \Omega_{M_k}) \geq \tilde{\theta} \text{osc}(f, T_k)$.

Refine each triangle $\tau \in \Omega_{M_k}$ by the newest vertex bisection to get $T_{k+1}$.

$k = k + 1$.

END WHILE

$T_N = T_k$.

END AMFEM

6.1. Convergence of AMFEM. We shall prove the algorithm AMFEM will terminate in finite steps by showing the reduction of the sum of the error and the error estimator.

We first summarize the main ingredients in the following lemma with the following short notation:

$e_k = \|\sigma - \sigma_k\|^2, E_k = \|\sigma_{k+1} - \sigma_k\|^2, o_k = \text{osc}^2(f, T_k)$, and $\eta_k = \eta^2(\sigma_k, E_k)$.

**Lemma 6.1.** One has the following inequalities

(6.3) $(1 - \delta) e_{k+1} \leq e_k - E_k + \frac{C_0}{\delta} o_k$, for any $\delta > 0$

(6.4) $\beta \eta_{k+1} \leq \beta(1 - \frac{1}{2}\theta) \eta_k + E_k,$

(6.5) $e_k \leq C_1 \eta_k + C_0 o_k.$

**Proof.** (6.3) is the quasi-orthogonality (3.4) established in Theorem 3.2 and Remark 3.3. (6.5) is the upper bound (5.3) in Theorem 5.7. We only need to prove (6.4). By the continuity of the error estimator (5.2), we have

(6.6) $\beta \eta^2(\sigma_{k+1}, E_{k+1}) \leq \beta \eta^2(\sigma_k, E_{k+1}) + E_k.$

Let $N_{k+1} = E_{k+1} \setminus E_k$ be the new edges in $T_{k+1}$ and $M_k \subseteq E_k$ be the refined edge in $T_k$.

It is obvious that $E_k \setminus M_k = E_{k+1} \setminus N_{k+1}$. For an edge $E \in N_{k+1}$, if it is an interior edge
of some triangle \( T \in \mathcal{T}_k \), then \( J_E(\sigma_k) = 0 \) since \( \sigma_k \) is a polynomial in \( T \). For other edges, it is at least half of some edge in \( \mathcal{M}_k \) and thus

\[
\eta^2(\sigma_k, \mathcal{N}_{k+1}) \leq \frac{1}{2} \eta^2(\sigma_k, \mathcal{M}_k).
\]

Since some edges are refined for the conformity of triangulation, \( \mathcal{M}_k \subseteq \mathcal{M}_k \). By the marking strategy (6.1), we have

\[
\eta^2(\sigma_k, \mathcal{M}_k) \geq \eta^2(\sigma_k, \mathcal{M}_k) \geq \theta \eta^2(\sigma_k, \mathcal{E}_k).
\]

Combining (6.7) and (6.8), we get

\[
\eta^2(\sigma_k, \mathcal{E}_{k+1}) = \eta^2(\sigma_k, \mathcal{N}_{k+1}) + \eta^2(\sigma_k, \mathcal{E}_{k+1} \setminus \mathcal{N}_{k+1})
\leq \frac{1}{2} \eta^2(\sigma_k, \mathcal{M}_k) + \eta^2(\sigma_k, \mathcal{E}_k \setminus \mathcal{M}_k)
\leq -\frac{1}{2} \eta^2(\sigma_k, \mathcal{M}_k) + \eta^2(\sigma_k, \mathcal{E}_k)
\leq (1 - \frac{1}{2}\theta)\eta^2(\sigma_k, \mathcal{E}_k).
\]

Substituting to (6.6) we then get (6.4). \( \square \)

**Theorem 6.2.** When

\[
0 < \delta < \min \{ \frac{\beta}{2C_1}, 1 \},
\]

there exists \( \alpha \in (0, 1) \) and \( C_\delta \) such that

\[
(1 - \delta)e_{k+1} + \beta \eta_{k+1} \leq \alpha \left[ (1 - \delta)e_k + \beta \eta_k \right] + C_\delta o_k.
\]

**Proof.** First (6.3) + (6.4) gives

\[
(1 - \delta)e_{k+1} + \beta \eta_{k+1} \leq e_k + \beta (1 - \frac{1}{2}\theta)\eta_k + \frac{C_0}{\delta} o_k.
\]

Then we separate \( e_H \) and use (6.4) to bound

\[
e_k = \alpha (1 - \delta)e_k + [1 - \alpha (1 - \delta)]e_k
\leq \alpha (1 - \delta)e_k + [1 - \alpha (1 - \delta)](C_1\eta_k + C_0 o_k).
\]

Therefore we obtain

\[
(1 - \delta)e_{k+1} + \beta \eta_{k+1} \leq \alpha \left[ (1 - \delta)e_k + \frac{[1 - \alpha (1 - \delta)]C_1}{\alpha} \eta_k \right] + C_\delta o_k.
\]

Now we choose \( \alpha \) such that

\[
\frac{1 - \alpha (1 - \delta)C_1}{\alpha} = \beta,
\]

i.e.

\[
\alpha = \frac{C_1 + (1 - \frac{1}{2}\theta)\beta}{C_1(1 - \delta) + \beta} = \frac{C_1 + \beta - \frac{1}{2}\theta\beta}{C_1 + \beta - C_1\delta}.
\]

By the requirement of \( \delta \) (6.9), we conclude \( \alpha \in (0, 1) \). \( \square \)

**Theorem 6.3.** Let \( \sigma_k \) be the solution obtained in the \( k \)-th loop in the algorithm AMFEM, then for any \( 0 < \delta < \min \{ \frac{\beta}{2C_1}, 1 \} \), there exist positive constants \( C_\delta \) and \( 0 < \gamma_\delta < 1 \) depending only on given data and the initial grid such that,

\[
(1 - \delta)\|\sigma - \sigma_k\|^2 + \beta \eta^2(\sigma_k, \mathcal{T}_k) \leq C_\delta \gamma_\delta^k,
\]

and thus the algorithm AMFEM will terminate in finite steps.
Proof. The proof is identical to that of Theorem 4.7 in [46] using (6.10).

6.2. Optimality of AMFEM. Let \( T_0 \) be an initial quasi-uniform triangulation with \( \#T_0 > 2 \) and \( P_N \) be the set of all triangulations \( T \) which is refined from \( T_0 \) and \( \#T \leq N \). For a given triangulation \( T \), the solution of the mixed finite element approximation of Poisson equation will be denoted by \( \sigma_T \). We define

\[
\mathcal{A}^s = \{ \sigma \in \Sigma : \|\sigma\|_{A^s} < \infty, \text{ with }\|\sigma\|_{A^s} = \sup_{N \geq \#T_0} \left( N^s \inf_{T \in P_N} \|\sigma - \sigma_T\| \right) \}.
\]

An adaptive finite element method realizes optimal convergence rates if whenever \( \sigma \in \mathcal{A}^s \), it produces approximation \( \sigma_N \) with respect to triangulations \( T_N \) elements such that

\[
\|\sigma - \sigma_N\| \leq C(\#T_N)^{-s}.
\]

For simplicity, we consider the following algorithm which separates the reduction of data oscillation and error:

1. \([T_H, f_H] = \text{APPROX}(f, T_0, \varepsilon/2)\)
2. \([\sigma_N, T_N] = \text{AMFEM}(T_H, f_H, \varepsilon/2, \theta, 0, 1)\)

The advantage of separating data and error is that in the second step, data oscillation is always zero since the input data \( f_H \) is piecewise polynomial in the initial grid \( T_H \) for AMFEM. In this case, we also list all ingredients needed for the optimality of adaptive procedure.

1. Orthogonality:
   \[
   \|\sigma - \sigma_{k+1}\|^2 = \|\sigma - \sigma_k\|^2 - \|\sigma_{k+1} - \sigma_k\|^2
   \]
2. Discrete upper bound:
   \[
   \|\sigma_{k+1} - \sigma_k\|^2 \leq C_1 \eta^2(\sigma_k, F_k) \text{ and } \#F_k \leq 3(\#T_{k+1} - \#T_k).
   \]
3. Lower bound:
   \[
   C_2 \eta^2(\sigma_k, E_k) \leq \|\sigma - \sigma_k\|^2.
   \]

Theorem 6.4. Let \([\sigma_N, T_N] = \text{AMFEM}(T_H, f_H, \varepsilon, \theta, 0, \mu)\), and \( \tilde{\sigma} = L^{-1} f_H \). If \( \tilde{\sigma} \in \mathcal{A}^s \), and \( 0 < \theta < C_2/C_1 \), then for any \( \varepsilon > 0 \), AMFEM will terminated in finite steps and

\[
(6.11) \quad \|\tilde{\sigma} - \sigma_N\| \leq \varepsilon, \quad \text{and} \quad \#T_N - \#T_0 \leq C\|\sigma\|_{A^s}^{1/s} \varepsilon^{-1/s}.
\]

Proof. It is identical to the proof of Theorem 5.3 in [55] using three ingredients listed above.

Theorem 6.5. For any \( f \in L^2(\Omega) \), a shape regular triangulation \( T_0 \) and \( \varepsilon > 0 \). Let \( \sigma = L^{-1} f \) and \([\sigma_N, T_N] = \text{AMFEM}(T_H, f_H, \varepsilon/2, 0, 1)\) where \([T_H, f_H] = \text{APPROX}(f, T_0, \varepsilon/2)\).

If \( \sigma \in \mathcal{A}^s \) and \( f \in \mathcal{A}^{s}_0 \), then

\[
\|\sigma - \sigma_N\| \leq C(\|\sigma\|_{A^s} + \|f\|_{A^s}) (\#T_N - \#T_0)^{-s}.
\]

Proof. Let \( \tilde{\sigma} = L^{-1} f_H \). By Theorem 4.1 and 2.2, we have

\[
(6.12) \quad \|\sigma - \tilde{\sigma}\| \leq \varepsilon/2, \quad \text{and} \quad \#T_H - \#T_0 \leq C\|f\|_{A^s}^{1/s} \varepsilon^{-1/s}.
\]

It is easy to show, by the definition of \( \mathcal{A}^s \), if \( \sigma \in \mathcal{A}^s \), then \( \tilde{\sigma} \in \mathcal{A}^s \) and

\[
\|\tilde{\sigma}\|_{A^s} \leq \|\sigma\|_{A^s} + \|f\|_{A^s}.
\]

We then apply Theorem 6.4 to \( \tilde{\sigma} \) to obtain

\[
(6.13) \quad \|\tilde{\sigma} - \sigma_N\| \leq \varepsilon/2 \quad \text{and} \quad \#T_N - \#T_H \leq C\|\tilde{\sigma}\|_{A^s}^{1/s} \varepsilon^{-1/s}.
\]
Combining (6.12) and (6.13) we get
\[ \| \sigma - \sigma_N \| \leq \| \sigma - \tilde{\sigma} \| + \| \tilde{\sigma} - \sigma_N \| \leq \varepsilon \]
and
\[ \varepsilon \leq C (\# T_N - \# T_0)^{-s} \left( \| \sigma \|_{A^s} + \| f \|_{A^s} \right). \]
The desired result then follows.

Acknowledgement. The authors would like to thank Dr. Guzman for the discussion on the simplification of the proof of the discrete stability result and Prof. Nochetto for the simplification of convergence analysis without using discrete lower bound.

REFERENCES


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT SAN DIEGO, LA JOLLA, CA 92093
E-mail address: lyc102@gmail.com

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT SAN DIEGO, LA JOLLA, CA 92093
E-mail address: mholst@math.ucsd.edu

THE SCHOOL OF MATHEMATICAL SCIENCE, PEKING UNIVERSITY, AND DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16801
E-mail address: xu@math.psu.edu