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 that is not divergence free 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.

CONTENTS

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 [8, 36, 48, 50, 12, 59, 51, 27, 29].

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

$$-\Delta u = f \quad \text{in } \Omega, \quad \text{and} \quad u = 0 \quad \text{on } \partial\Omega, \tag{1.1}$$

posed on a polygonal and simply connected domain $\Omega \subset \mathbb{R}^2$. In many applications ([24]) the variable $\boldsymbol{\sigma} = -\nabla u$ is of interest and it is therefore convenient to use mixed finite element methods, such as the Raviart-Thomas mixed method [53] and Brezzi-Douglas-Marini mixed method [23]. 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 $\boldsymbol{\sigma}$ in an optimal way.

Our main result is the following optimal convergence of our algorithms **AMFEM** and its variant. Let $\boldsymbol{\sigma}_N$ be the approximation of $\boldsymbol{\sigma}$ based on the triangulation $\mathcal{T}_N$ obtained in **AMFEM**. If $\boldsymbol{\sigma} \in \mathcal{A}^s$ and $f \in \mathcal{A}_o^s$, then

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_N\| \le C(\|\boldsymbol{\sigma}\|_{\mathcal{A}^s} + \|f\|_{\mathcal{A}_o^s})(\#\mathcal{T}_N - \#\mathcal{T}_0)^{-s}, \tag{1.2}$$

where $(\mathcal{A}^s, \|\cdot\|_{\mathcal{A}^s})$ and $(\mathcal{A}_o^s, \|\cdot\|_{\mathcal{A}_o^s})$ are approximation spaces as in [12]. The index s is used to characterize the best possible approximation rate of $\boldsymbol{\sigma}$, which depends on the regularity of the solution and data, and the order of elements. For example when $f \in L^2(\Omega)$ and $\boldsymbol{\sigma} \in \boldsymbol{W}^{1,1}(\Omega)$, we can achieve the optimal convergence rate $s = 1/2$ for the lowest order Raviart-Thomas finite element space. We refer to [13] for the characterization of $\mathcal{A}^s$ in terms of Besov spaces and to [9, 10, 35, 34] 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 $\boldsymbol{\sigma}$.

For the analysis of the convergence of adaptive procedure, we follow the new approach by Cascon, Kreuzer, Nochetto and Siebert [27], and for the optimality we mainly use the simplified case in Stevenson's work [59]. 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 [26], 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 *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 *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 primal variational formulation: for a given $f \in L^2(\Omega)$, find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv \qquad \forall v \in H_0^1(\Omega), \tag{2.1}$$

and then finding an approximation by solving (2.1) in finite-dimensional subspaces of $H_0^1(\Omega)$. In many applications ([24]) the variable $\boldsymbol{\sigma} = -\nabla u$ 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:

$$\boldsymbol{\sigma} + \nabla u = 0, \ \operatorname{div} \boldsymbol{\sigma} = f \ \text{in} \ \Omega, \quad \text{and} \quad u = 0 \ \text{on} \ \partial\Omega. \tag{2.2}$$

Let

$$\boldsymbol{\Sigma} = \boldsymbol{H}(\operatorname{div};\Omega) := \{\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega) : \operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)\}, \ \text{and} \ U = L^2(\Omega).$$

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

$$\|\boldsymbol{\tau}\|_{H(\operatorname{div})} = (\|\boldsymbol{\tau}\|^2 + \|\operatorname{div} \boldsymbol{\tau}\|^2)^{1/2}, \quad \forall \tau \in \boldsymbol{\Sigma}.$$

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

$$(\boldsymbol{\sigma}, \boldsymbol{\tau}) - (\operatorname{div} \boldsymbol{\tau}, u) = 0 \qquad\qquad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \tag{2.3}$$

$$(\operatorname{div} \boldsymbol{\sigma}, v) = (f, v) \qquad\qquad \forall v \in U, \tag{2.4}$$

where $(\cdot, \cdot)$ is the inner product for $L^2(\Omega)$ or $\boldsymbol{L}^2(\Omega)$. 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 $(\boldsymbol{\sigma}, u)$ to (2.3)-(2.4) follows from the so-called inf-sup condition which can be easily established for this model problem [24].

Given a shape regular and conforming (in the sense of [30]) triangulation $\mathcal{T}_H$ of Ω, 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 $(\boldsymbol{\sigma}_H, u_H) \in \Sigma_H \times U_H$ such that

$$(\boldsymbol{\sigma}_H, \boldsymbol{\tau}_H) - (\operatorname{div} \boldsymbol{\tau}_H, u_H) = 0 \qquad \forall \, \boldsymbol{\tau}_H \in \Sigma_H \qquad (2.5)$$

$$(\operatorname{div} \boldsymbol{\sigma}_H, v_H) = (f_H, v_H) \qquad \forall \, v_H \in U_H. \qquad (2.6)$$

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 [53] and Brezzi-Douglas-Marini spaces [23]. Recently it has been shown that such stable finite element spaces can be constructed in an elegant way using differential complex theory [16, 41, 2, 5].

The Raviart-Thomas spaces [53] are defined for an integer $p \geq 0$ by

$$RT_H = \boldsymbol{\Sigma}_H^p \times U_H^p, \text{ where}$$

$$\boldsymbol{\Sigma}_H^p(\mathcal{T}_H) := \{\boldsymbol{\tau} \in H(\operatorname{div}; \Omega) \, : \, \boldsymbol{\tau}|_T \in \boldsymbol{P}_p(T) + \boldsymbol{x}\mathcal{P}_p(T), \; \forall T \in \mathcal{T}_H\},$$

$$\text{and} \quad U_H^p(\mathcal{T}_H) := \{v \in L^2(\Omega) \, : \, v|_T \in \mathcal{P}_p(T), \; \forall T \in \mathcal{T}_H\},$$

and where $\mathcal{P}_p(T)$ denotes the space of polynomials on T of degree at most p.

The Brezzi-Douglas-Marini spaces [23] are defined for an integer $p \geq 1$ by

$$BDM_H = \boldsymbol{\Sigma}_H^p \times U_H^p, \text{ where}$$

$$\boldsymbol{\Sigma}_H^p(\mathcal{T}_H) := \{\boldsymbol{\tau} \in H(\operatorname{div}; \Omega) \, : \, \boldsymbol{\tau}|_T \in \boldsymbol{P}_p(T), \; \forall T \in \mathcal{T}_H\},$$

$$\text{and} \quad U_H^p(\mathcal{T}_H) := \{v \in L^2(\Omega) \, : \, v|_T \in \mathcal{P}_{p-1}(T), \; \forall T \in \mathcal{T}_H\}.$$

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 (Σ_H, U_H) to denote the pair in RT_H or BDM_H. The discrete problem posed on (Σ_H, U_H) will satisfy the discrete inf-sup condition [24] from which the existence and uniqueness of the finite element approximation $(\boldsymbol{\sigma}_H, u_H)$ 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

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

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

2.2. **Adaptive methods through local refinement.** Let $\boldsymbol{\sigma} = \mathcal{L}^{-1}f$ and $\boldsymbol{\sigma}_H = \mathcal{L}_H^{-1}f_H$. We are mostly interested in the control of the error $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|$ 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 $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|_{H(\operatorname{div})}$, we comment that, by (2.4) and (2.6), $\|\operatorname{div} \boldsymbol{\sigma} - \operatorname{div} \boldsymbol{\sigma}_H\| = \|f - f_H\|$ can be approximated efficiently without solving equations and also may dominate the error $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|_{H(\operatorname{div})}$; see Remark 3.4 in [45].

The rate of the error $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|$ for $\boldsymbol{\sigma}_H \in \boldsymbol{\Sigma}_H^p(\mathcal{T}_H)$ depends on the regularity of the function being approximated and the regularity of the mesh. If $\boldsymbol{\sigma} \in \boldsymbol{H}^{p+1}(\Omega)$ and $\mathcal{T}_H$ is

quasi-uniform with mesh size $H = \max_{T \in \mathcal{T}_H} \operatorname{diam}(T)$, then the following convergence result of optimal order is well known [24]

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

The regularity result $\boldsymbol{\sigma} \in \boldsymbol{H}^{p+1}(\Omega)$, however, may not be true in many applications, especially for concave domains Ω. 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, $\#\mathcal{T}$, 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 $\mathcal{T}_0$, we shall generate a sequence of nested conforming triangulations $\mathcal{T}_k$ using the following loop

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

More precisely, to get $\mathcal{T}_{k+1}$ from $\mathcal{T}_k$ we first solve (2.5)-(2.6) to get $\boldsymbol{\sigma}_k$ on $\mathcal{T}_k$. The error is estimated using $\boldsymbol{\sigma}_k$ 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 [30].

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 [17, 18, 20, 21, 40, 52, 56].

The *a posteriori* error estimators are essential part of the **ESTIMATE** step. Given a shape regular triangulation $\mathcal{T}_H$, let $\mathcal{E}_H$ denote the edges of $\mathcal{T}_H$. In this paper, we shall use edge-wise error estimator η_E for each edge $E \in \mathcal{E}_H$. See Section 4 for details.

The local error estimator η_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 [36] 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 [11], longest edge refinement [55, 54], and newest vertex bisection [58, 46, 47]. 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 [46, 61, 12, 28] for details of newest vertex bisection and only list two important properties below.

Let $\mathcal{T}_k$ be a conforming triangulation refined from a shape regular triangulation $\mathcal{T}_0$ using the new vertex bisection and let $\mathcal{M}$ be the collection of all marked triangles going from $\mathcal{T}_0$ to $\mathcal{T}_k$. Then

(1) $\{\mathcal{T}_k\}$ is shape regular and the shape regularity only depends on $\mathcal{T}_0$;
(2) $\#\mathcal{T}_k \leq \#\mathcal{T}_0 + C\#\mathcal{M}$.

Recently Stevenson [60] showed 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 extended to general space dimensions.

2.3. **Approximation of the data.** We shall introduce the concept of data oscillation which is firstly introduced in [48], 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 simplify the notation, we may drop the subscript if it is clear from the context. For a triangulation $\mathcal{T}_H$ of Ω and a function $f \in L^2(\Omega)$, we define a triangulation dependent norm

$$\|H f\|_{0,\mathcal{T}_H} := \Big(\sum_{T \in \mathcal{T}_H} H_T^2 \|f\|_{0,T}^2 \Big)^{1/2}.$$

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

$$\operatorname{osc}(f, \mathcal{T}_H) := \|H(f - f_H)\|_{0,\mathcal{T}_H}.$$

Let $\mathcal{P}_N$ denote the set of triangulations constructed from an initial triangulation $\mathcal{T}_0$ by the newest vertex bisection method with at most N triangles. We define

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

where N_0 is a fixed integer representing the number of triangles in $\mathcal{T}_0$. We will recall a result of Binev, Dahmen and DeVore [12] which shows that the approximation of data can be done in an optimal way. The proof can be found at [12]; See also [14].

Theorem 2.2 (Binev, Dahmen and DeVore). *Given a tolerance ε, an $f \in L^2(\Omega)$ and a shape regular triangulation $\mathcal{T}_0$, there exists an algorithm*

$$\mathcal{T}_H = \textbf{APPROX}(f, \mathcal{T}_0, \varepsilon)$$

such that

$$\operatorname{osc}(f, \mathcal{T}_H) \leq \varepsilon, \quad and \quad \#\mathcal{T}_H - \#\mathcal{T}_0 \leq C\|f\|_{\mathcal{A}_o^{1/s}}^{1/s} \varepsilon^{-1/s}.$$

3. QUASI-ORTHOGONALITY

Unlike the primal formulation of Poisson equation, $\boldsymbol{\sigma}_H$ is not the L^2-orthogonal projection of $\boldsymbol{\sigma}$ from Σ to Σ_H. Indeed the solution $(\boldsymbol{\sigma}, u)$ of (2.3)-(2.4) is the saddle point of the following energy

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

Namely

$$E(\boldsymbol{\sigma}, u) = \inf_{\boldsymbol{\sigma} \in \boldsymbol{H}(\operatorname{div};\Omega)} \sup_{v \in L^2(\Omega)} E(\boldsymbol{\tau}, v).$$

Similar result holds for the discrete solutions $(\boldsymbol{\sigma}_H, u_H)$. 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 $\boldsymbol{\sigma} - \boldsymbol{\sigma}_H$ is orthogonal to the divergence free subspace of Σ_H to prove a quasi-orthogonality result. In the sequel we shall consider two conforming triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$ which are nested in the sense that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. 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 $\mathcal{T}_h$ and $\mathcal{T}_H$, let*
$$(\boldsymbol{\sigma}, u) = \mathcal{L}^{-1} f, (\boldsymbol{\sigma}_h, u_h) = \mathcal{L}_h^{-1} f_h, (\tilde{\boldsymbol{\sigma}}_h, \tilde{u}_h) = \mathcal{L}_h^{-1} f_H, \text{ and } (\boldsymbol{\sigma}_H, u_H) = \mathcal{L}_H^{-1} f_H.$$
Then
$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H) = 0. \tag{3.1}$$

Proof. Since $\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H \in \Sigma_h$, by (2.5)-(2.6), we have
$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H) = (u - u_h, \operatorname{div}(\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H)) = (u - u_h, f_H - f_H) = 0.$$
$\square$

To prove quasi-orthogonality, we need the following discrete stability result
$$\|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\| \le \sqrt{C_0} \operatorname{osc}(f_h, \mathcal{T}_H), \tag{3.2}$$
where the constant C_0 depends only on the shape regularity of $\mathcal{T}_H$. 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 triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$, let $\boldsymbol{\sigma} = \mathcal{L}^{-1} f$, $\boldsymbol{\sigma}_h = \mathcal{L}_h^{-1} f_h$, and $\boldsymbol{\sigma}_H = \mathcal{L}_H^{-1} f_H$. Then*
$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H) \le \sqrt{C_0} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| \operatorname{osc}(f_h, \mathcal{T}_H), \tag{3.3}$$
Thus for any $\delta > 0$,
$$(1 - \delta)\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2 - \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 + \frac{C_0}{\delta} \operatorname{osc}^2(f_h, \mathcal{T}_H), \tag{3.4}$$
and in particular when $\operatorname{osc}(f_h, \mathcal{T}_H) = 0$,
$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2 - \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2. \tag{3.5}$$

Proof. Let us introduce an intermediate solution $\tilde{\boldsymbol{\sigma}}_h = \mathcal{L}_h^{-1} f_H$. By Lemma 3.1, $(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H) = 0$. Thus
$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H) = (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h) \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| \|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|.$$
(3.3) then follows from the inequality (3.2).

By the trivial identity $\boldsymbol{\sigma} - \boldsymbol{\sigma}_H = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h + \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H$, we have
$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2 = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 + 2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H)$$
When $\operatorname{osc}(f_h, \mathcal{T}_H) = 0$, by (3.3), $(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H) = 0$ and thus (3.5) follows. In general, we use
$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2 &= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 + 2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H) \\
&\ge \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 - 2\sqrt{C_0} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| \operatorname{osc}(f, \mathcal{T}_H) \\
&\ge \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 + (1 - \delta)\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 - \frac{C_0}{\delta} \operatorname{osc}^2(f_h, \mathcal{T}_H),
\end{aligned}$$
to prove (3.4). In the last step, we have used the inequality
$$2ab \le \delta a^2 + \frac{1}{\delta} b^2, \quad \text{for any } \delta > 0.$$
$\square$

A similar quasi-orthogonality result was obtained by Carstensen and Hoppe [26] 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 [31, 32, 33] for discussion on this relation. In contrast the approach we used here is more straight-forward.

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

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

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

4. DISCRETE STABILITY FOR PERTURBATION OF DATA

In this section, we shall prove the discrete stability result. We begin with a stability result in the continuous case. Let $u \in H_0^1(\Omega)$ be the solution of the primal weak formulation (2.1) of Poisson equation. Then $(-\nabla u, u)$ is the solution to the dual weak formulation (2.3)-(2.4). The stability result $\|\nabla u\| \le \|f\|_{-1}$ 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 $\mathcal{T}_H$ of Ω and $f \in L^2(\Omega)$, let $(\boldsymbol{\sigma}, u) = \mathcal{L}^{-1}f$ and $(\tilde{\boldsymbol{\sigma}}, \tilde{u}) = \mathcal{L}^{-1}f_H$, respectively. Then there exists a constant C_0 depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$\|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\| \le \sqrt{C_0}\operatorname{osc}(f, \mathcal{T}_H). \tag{4.1}$$

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

$$\|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|^2 = (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}, \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}) = (\operatorname{div}(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}), u - \tilde{u}) = (f - f_H, u - \tilde{u}).$$

Let v be the solution of primal weak formulation of Poisson equation with data $f - f_H$. Then $v = u - \tilde{u}$ and $-\nabla v = \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}$. Recall that $Q_H : L^2(\Omega) \to U_H$ is the L^2 projection into discontinuous polynomial spaces. So for each triangle $T \in \mathcal{T}_H$, $(f - f_H, v_H)_T = 0$ for any $v_H \in \mathcal{P}_p(T)$. Therefore

$$
\begin{aligned}
\|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|^2 &= (f - f_H, v) \\
&= \sum_{T \in \mathcal{T}_H} (f - f_H, v - Q_H v)_T \\
&\le \sqrt{C_0} \sum_{T \in \mathcal{T}_H} \|H(f - f_H)\|_{0,T}\|\nabla v\|_{0,T} \\
&\le \sqrt{C_0}\Big(\sum_{T \in \mathcal{T}_H} \|H(f - f_H)\|_{0,T}^2\Big)^{1/2}\|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|.
\end{aligned}
$$

In the second step, we have used the error estimate

$$\|v - Q_H v\|_{0,T} \le \sqrt{C_0}H_T\|\nabla v\|_{0,T},$$

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

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

$$\|u - Q_H u\|_{0,T} \le \sqrt{C_0}H_T\|\nabla u\|_{0,T} = \sqrt{C_0}H_T\|\boldsymbol{\sigma}\|_T,$$

for $u \in H_0^1(\Omega)$ and $\boldsymbol{\sigma} = -\nabla u$. The main difficulty in the discrete case is that $u_h \in U_h \not\subseteq H_0^1(\Omega)$. However we still have a similar localized error estimate for $u_h - Q_H u_h$.

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

$$\|u_h - Q_H u_h\|_{0,T} \le \sqrt{C_0} H_T \|\boldsymbol{\sigma}_h\|_{0,T}. \tag{4.2}$$

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 $\mathcal{T}_h$ and $\mathcal{T}_H$ be two nested conforming triangulations. Let $\tilde{\boldsymbol{\sigma}}_h = \mathcal{L}_h^{-1} f_H$ and $\boldsymbol{\sigma}_h = \mathcal{L}_h^{-1} f_h$. Then there exists a constant C_0, depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$\|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\| \le \sqrt{C_0} \operatorname{osc}(f_h, \mathcal{T}_H). \tag{4.3}$$

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

$$(\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) = (u_h - \tilde{u}_h, \operatorname{div} \boldsymbol{\tau}_h), \qquad \forall \boldsymbol{\tau}_h \in \Sigma_h \tag{4.4}$$

$$(\operatorname{div}(\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h), v_h) = (f_h - f_H, v_h), \qquad \forall v_h \in U_h. \tag{4.5}$$

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

$$\|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|^2 = (u_h - \tilde{u}_h, \operatorname{div}(\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h)) = (v_h, f_h - f_H) = (v_h - Q_H v_h, f_h - f_H).$$

In the third step, we use the fact $f_H = Q_H f = Q_H f_h$ since $\mathcal{T}_h$ and $\mathcal{T}_H$ are nested. Thanks to (4.2), we have

$$\begin{aligned}
\|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|^2 &= \sum_{T \in \mathcal{T}_H} (v_h - Q_H v_h, f_h - f_H)_T \\
&\le \sqrt{C_0} \sum_{T \in \mathcal{T}_H} H_T \|f_h - f_H\|_{0,T} \|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|_{0,T} \\
&\le \sqrt{C_0} \left(\sum_{T \in \mathcal{T}_H} H_T^2 \|f_h - f_H\|_T^2 \right)^{1/2} \|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|.
\end{aligned}$$

Canceling one $\|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|$, we get the desired result. $\qquad\square$

In the rest of this section, we shall prove Lemma 4.2. It is a modification of arguments in [4] from quasi-uniform grids to adaptive grids. The first ingredient is the existence of a continuous right inverse of the divergence as an operator from $\boldsymbol{H}_0^1(\Omega)$ into the space $L_0^2(\Omega) := \{v \in L^2(\Omega) : \int_\Omega v = 0.\}$

Lemma 4.4. *Given a function $f \in L_0^2(\Omega)$, there exists a function $\boldsymbol{\tau} \in \boldsymbol{H}_0^1(\Omega)$ such that*

$$\operatorname{div} \boldsymbol{\tau} = f \quad \text{and} \quad \|\boldsymbol{\tau}\|_1 \le C \|f\|.$$

The proof of this lemma for smooth or convex domains Ω is pretty easy. One can solve the Poisson equation with Neumann boundary condition

$$\Delta \phi = f \text{ in } \Omega, \quad \frac{\partial \phi}{\partial \boldsymbol{n}} = 0 \text{ on } \partial\Omega.$$

The condition $f \in L_0^2(\Omega)$ ensures the existence of the solution. Then we let $\boldsymbol{\tau} = \operatorname{grad} \phi$ and modify the tangent component of $\boldsymbol{\tau}$ to be zero [22]. See also [7, 37] for a detailed proof on non-convex and general Lipschitz domains.

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

Lemma 4.5. *There exists an interpolation operator $\Pi_h : \boldsymbol{H}^1(T) \to \Sigma_h$ such that*

    (1) $Q_h \operatorname{div} \boldsymbol{\tau} = \operatorname{div} \Pi_h \boldsymbol{\tau}, \quad \forall \boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega)$;

(2) there exists a constant C depending only on the shape regularity of $\mathcal{T}_h$ such that

$$\|\boldsymbol{\tau} - \Pi_h \boldsymbol{\tau}\|_T \le Ch_T \|\boldsymbol{\tau}\|_{1,T}, \quad \forall T \in \mathcal{T}_h, \forall \boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega);$$

(3) for any $T \in \mathcal{T}_h$ if $\boldsymbol{\tau} \in H_0^1(T)$, then $\Pi_h \boldsymbol{\tau}|_{\partial T} = 0$.

For the detailed construction of such interpolation operator and proof of these properties, we refer to [41] and [6].

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 $T \in \mathcal{T}_H$, by the definition of L^2 projection Q_H, we have, $\int_T (Q_h - Q_H)u_h = 0$ i.e. $(Q_h - Q_H)u_h \in L_0^2(T)$. We thus can apply Lemma 4.4 to find a function $\boldsymbol{\tau} \in \boldsymbol{H}_0^1(T)$ such that

$$\operatorname{div} \boldsymbol{\tau} = (Q_h - Q_H)u_h, \text{ in } T \quad \text{and} \quad \|\boldsymbol{\tau}\|_{1,T} \le C\|(Q_h - Q_H)u_h\|_{0,T}.$$

We extend $\boldsymbol{\tau}$ to $\boldsymbol{H}^1(\Omega)$ by zero. Note that

$$(\Pi_h - \Pi_H)\boldsymbol{\tau} \in \Sigma_h, \text{ and } \operatorname{supp}(\Pi_h - \Pi_H)\boldsymbol{\tau} \subseteq T. \tag{4.6}$$

With such $\boldsymbol{\tau}$, we have

$$\|(Q_h - Q_H)u_h\|_{0,T}^2 = ((Q_h - Q_H)u_h, \operatorname{div} \boldsymbol{\tau})_T = (u_h, (Q_h - Q_H)\operatorname{div} \boldsymbol{\tau})_T.$$

Then using the commuting property (Lemma 4.5 (1)) and the locality of $\boldsymbol{\tau}$, we have

$$(u_h, (Q_h - Q_H)\operatorname{div} \boldsymbol{\tau})_T = (u_h, (Q_h - Q_H)\operatorname{div} \boldsymbol{\tau})_\Omega = (u_h, \operatorname{div}(\Pi_h - \Pi_H)\boldsymbol{\tau})_\Omega.$$

Now we shall use the fact $(\boldsymbol{\sigma}_h, u_h)$ is the solution of (2.3) and (2.4) and, again, the locality of $\boldsymbol{\tau}$ to get

$$(u_h, \operatorname{div}(\Pi_h - \Pi_H)\boldsymbol{\tau})_\Omega = (\boldsymbol{\sigma}_h, (\Pi_h - \Pi_H)\boldsymbol{\tau})_\Omega = (\boldsymbol{\sigma}_h, (\Pi_h - \Pi_H)\boldsymbol{\tau})_T.$$

Using the approximation property of Π_h (Lemma 4.5 (2)), we get

$$\begin{aligned} (\boldsymbol{\sigma}_h, (\Pi_h - \Pi_H)\boldsymbol{\tau})_T &\le \|\boldsymbol{\sigma}_h\|_{0,T}(\|\boldsymbol{\tau} - \Pi_h \boldsymbol{\tau}\|_{0,T} + \|\boldsymbol{\tau} - \Pi_H \boldsymbol{\tau}\|_{0,T}) \\ &\le CH_T \|\boldsymbol{\sigma}_h\|_{0,T} \|\boldsymbol{\tau}\|_{1,T}. \end{aligned}$$

So we have

$$\|(Q_h - Q_H)u_h\|_{0,T}^2 \le CH_T \|\boldsymbol{\sigma}_h\|_{0,T} \|\boldsymbol{\tau}\|_{1,T} \le CH_T \|\boldsymbol{\sigma}_h\|_T \|(Q_h - Q_H)u_h\|_{0,T}.$$

Canceling one $\|(Q_h - Q_H)u_h\|_T$, we obtain the desired result. $\qquad\square$

5. A Posterior 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 [25, 62, 43, 39, 44, 45]. 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 $\boldsymbol{t}_E$ for E. We denote the patch of E consisting of triangles sharing E by Ω_E.

Definition 5.1. *Given a triangulation $\mathcal{T}_H$, for an $E \in \mathcal{E}_H$ and $E \notin \partial\Omega$, let $\Omega_E = T \cup \tilde{T}$. For any $\boldsymbol{\sigma}_H \in \Sigma_H$, we define the jump of $\boldsymbol{\sigma}_H$ across edge E as*

$$J_E(\boldsymbol{\sigma}_H) = [\boldsymbol{\sigma}_H \cdot \boldsymbol{t}_E] := \boldsymbol{\sigma}_H|_T \cdot \boldsymbol{t}_E - \boldsymbol{\sigma}_H|_{\tilde{T}} \cdot \boldsymbol{t}_E. \tag{5.1}$$

If $E \in \mathcal{E}_H \cap \partial\Omega$, we define $J_E(\boldsymbol{\sigma}_H) = \boldsymbol{\sigma}_H \cdot \boldsymbol{t}_E$. The edge error estimator is defined as

$$\eta_E^2(\boldsymbol{\sigma}_H) = \|H \operatorname{rot} \boldsymbol{\sigma}_H\|_{0,\Omega_E}^2 + \|H^{1/2} J_E(\boldsymbol{\sigma}_H)\|_{0,E}^2.$$

For a subset $\mathcal{F}_H \subseteq \mathcal{E}_H$, we define

$$\eta^2(\boldsymbol{\sigma}_H, \mathcal{F}_H) := \sum_{E \in \mathcal{F}_H} \eta_E^2(\boldsymbol{\sigma}_H).$$

The error estimator $\eta_E(\boldsymbol{\sigma}_H)$ is continuous with respect to $\boldsymbol{\sigma}_H$ 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 $\mathcal{T}_H$, let $\boldsymbol{\sigma}_H, \boldsymbol{\tau}_H \in \Sigma_H$. There exists constant β such that*

$$\beta \left| \eta^2(\boldsymbol{\sigma}_H, \mathcal{E}_H) - \eta^2(\boldsymbol{\tau}_H, \mathcal{E}_H) \right| \leq \|\boldsymbol{\sigma}_H - \boldsymbol{\tau}_H\|^2. \tag{5.2}$$

Proof. It can be easily proved by the triangle inequality and inverse inequality. $\square$

We shall recall Alonso's results below and prove a discrete upper bound later. Since the data f is not included in the definition of our error estimator η_E, the upper bound contains an additional data oscillation term which is different from the standard one in [61].

Theorem 5.3 (Upper bound). *Given an $f \in L^2(\Omega)$ and a shape regular triangulation $\mathcal{T}_H$, let $\boldsymbol{\sigma} = \mathcal{L}^{-1} f$ and $\boldsymbol{\sigma}_H = \mathcal{L}_H^{-1} f_H$. There exist constants C_0 and C_1 depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2 \leq C_1 \eta^2(\boldsymbol{\sigma}_H, \mathcal{E}_H) + C_0 \mathrm{osc}^2(f, \mathcal{T}_H). \tag{5.3}$$

Theorem 5.4 (Lower bound). *Given an $f \in L^2(\Omega)$ and a shape regular triangulation $\mathcal{T}_H$, let $\boldsymbol{\sigma} = \mathcal{L}^{-1} f$ and $\boldsymbol{\sigma}_H = \mathcal{L}_H^{-1} f_H$. There exists constant C_2 depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$C_2 \eta^2(\boldsymbol{\sigma}_H, \mathcal{E}_H) \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|^2, \tag{5.4}$$

for Raviart-Thomas spaces.

For Brezzi-Douglas-Marini spaces, (5.4) holds when $\mathrm{osc}(f, \mathcal{T}_H) = 0$.

When $\mathrm{osc}(f, \mathcal{T}_H) = 0$, (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 $\mathcal{T}_h$, let

$$S_h^p = \{\psi_h \in \mathcal{C}(\overline{\Omega}) : \psi_h|_T \in \mathcal{P}_p(T), \ \forall T \in \mathcal{T}_h\}$$

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 of $\mathrm{div} : \Sigma_h \mapsto U_h$.

Definition 5.5. *We define $\mathrm{grad}_h : U_h \mapsto (\Sigma_h)^*$ by*

$$(\mathrm{grad}_h v_h, \boldsymbol{\tau}_h) = -(v_h, \mathrm{div}\, \boldsymbol{\tau}_h), \quad \forall \boldsymbol{\tau}_h \in \Sigma_h.$$

We emphasis that grad_h is not simply the restriction of grad to U_h since U_h is not a subspace of $H^1(\Omega)$. The following discrete Helmholtz decomposition is well known in the literature; See, for example, [38, 19, 5, 15].

Theorem 5.6 (Discrete Helmholtz Decomposition in $\mathbb{R}^2$). *Given a triangulation $\mathcal{T}_h$, for p-th order Raviart-Thomas finite element spaces (Σ_h^p, U_h^p), we have the following orthogonal (with respect to L^2 inner product) decomposition*

$$\Sigma_h^p = \mathrm{curl}(S_h^{p+1}) \oplus \mathrm{grad}_h(U_h^p).$$

For Brezzi-Douglas-Marini finite element spaces (Σ_h^p, U_h^p), we have the following orthogonal (with respect to L^2 inner product) decomposition

$$\Sigma_h^p = \mathrm{curl}(S_h^{p+1}) \oplus \mathrm{grad}_h (U_h^{p-1}).$$

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

Theorem 5.7. *Let $\mathcal{T}_h$ and $\mathcal{T}_H$ be two nested conforming triangulations. Let $\boldsymbol{\sigma}_h = \mathcal{L}_h^{-1} f_h$ and $\boldsymbol{\sigma}_H = \mathcal{L}_H^{-1} f_H$, and let $\mathcal{F}_H = \{E \in \mathcal{E}_H : E \notin \mathcal{E}_h\}$. Then there exist constants depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 \leq C_1 \eta^2(\boldsymbol{\sigma}_H, \mathcal{F}_H) + C_0 \mathrm{osc}^2(f_h, \mathcal{T}_H) \tag{5.5}$$

and

$$\#\mathcal{F}_H \leq 3(\#\mathcal{T}_h - \#\mathcal{T}_H). \tag{5.6}$$

Proof. The inequality (5.6) follows from

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

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

$$\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H = \mathrm{grad}_h \phi_h + \mathrm{curl}\, \psi_h,$$

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.

$$\|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|^2 = \|\mathrm{grad}_h \phi_h\|^2 + \|\mathrm{curl}\, \psi_h\|^2. \tag{5.7}$$

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

$$|\psi_h|_1 \leq \|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|. \tag{5.8}$$

Since

$$(\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H, \mathrm{grad}_h v_h) = (\mathrm{div}\,(\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H), v_h) = (f_H - f_H, v_h - v_H) = 0,$$

we have

$$\|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|^2 = (\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H, \mathrm{grad}_h \phi_h) + (\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H, \mathrm{curl}\, \psi_h) = (\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H, \mathrm{curl}\, \psi_h).$$

Since $\mathrm{div}\, \mathrm{curl}\, \psi = 0$, $(\tilde{\boldsymbol{\sigma}}_h, \mathrm{curl}\, \psi_h) = 0$ and $(\boldsymbol{\sigma}_H, \mathrm{curl}\, \psi_H) = 0$ for any $\psi_H \in S_H^{p+1}$.

Choosing $\psi_H = \mathcal{I}_H \psi_h$ using some local quasi-interpolation, for example the Scott-Zhang quasi-interpolation [57], $\mathcal{I}_H : S_h^{p+1} \mapsto S_H^{p+1}$ such that

$$\|\psi_h - \mathcal{I}_H \psi_h\|_{0,E} \leq C H_E^{1/2} |\psi_h|_{1,\Omega_E} \quad \text{and} \quad \|\psi_h - \mathcal{I}_H \psi_h\|_{0,T} \leq C H_T |\psi_h|_{1,\Omega_T},$$

where $\Omega_T = \{T_H \subset \mathcal{T}_H : T_H \cap T \neq \emptyset\}$. Furthermore the quasi-interpolation $\mathcal{I}_H$ is local in the sense that if $T \in \mathcal{T}_H \cap \mathcal{T}_h$ or $E \in \mathcal{E}_H \cap \mathcal{E}_h$ (i.e. T or E is not refined), then $(\psi_h - \mathcal{I}_H \psi_h)|_T = 0$ or $(\psi_h - \mathcal{I}_H \psi_h)|_E = 0$, respectively. With such choice of $\mathcal{F}_H$ and ψ_H, we have

$$\begin{aligned}
\|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|^2 &= (\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H, \mathrm{curl}\, \psi_h) = (-\boldsymbol{\sigma}_H, \mathrm{curl}\,(\psi_h - \psi_H)) \\
&= \sum_{T \in \mathcal{T}_H} \left[\sum_{E \in \partial T} (\boldsymbol{\sigma}_H \cdot \boldsymbol{t}_E, \psi_h - \psi_H)_E + (\mathrm{rot}\, \boldsymbol{\sigma}_H, \psi_h - \psi_H)_T \right] \\
&\leq \sum_{E \in \mathcal{E}_H} [\boldsymbol{\sigma}_H] \|\psi_h - \psi_H\|_{0,E} + \sum_{T \in \mathcal{T}_H} \|\mathrm{rot}\, \boldsymbol{\sigma}_H\| \|\psi_h - \psi_H\|_{0,T}, \\
&\leq C \eta(\boldsymbol{\sigma}_H, \mathcal{F}_H) |\psi_h|_1 \leq C_1 \eta(\boldsymbol{\sigma}_H, \mathcal{F}_H) \|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|.
\end{aligned}$$

Canceling one $\|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|$, we get

$$\|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\| \leq C\eta(\boldsymbol{\sigma}_H, \mathcal{F}_H). \tag{5.9}$$

Now we write $\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H = \boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h + \tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H$ and note that

$$(\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h, \tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H) = (u_h - \tilde{u}_h, \operatorname{div}(\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H)) = (u_h - \tilde{u}_h, f_H - f_H) = 0.$$

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

$$\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|^2 = \|\tilde{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_H\|^2 + \|\boldsymbol{\sigma}_h - \tilde{\boldsymbol{\sigma}}_h\|^2 \leq C_1\eta^2(\boldsymbol{\sigma}_H, \mathcal{F}_H) + C_0\operatorname{osc}^2(f_h, \mathcal{T}_H).$$

$\square$

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 [48, 59, 36, 49, 50]. For the completeness we include them here and prove some important technical results.

We first present our algorithms. It mainly follows from the algorithm proposed in [50]. The difference is that we do not impose an interior point property in the refinement step.

Let $\mathcal{T}_0$ be a initial shape regular triangulation, a right hand side $f \in L^2(\Omega)$, a tolerance ε, 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 $\mathcal{M}_k$, we denote by $\Omega_{\mathcal{M}_k} = \cup_{E \in \mathcal{M}_k} \Omega_E$.

$[\mathcal{T}_N, \boldsymbol{\sigma}_N]$=**AMFEM**$(\mathcal{T}_0, f, \varepsilon, \theta, \tilde{\theta}, \mu)$
   $\eta = \varepsilon, k = 0$
  WHILE $\eta \geq \varepsilon$, **DO**
    Solve (2.5)-(2.6) on $\mathcal{T}_k$ to get the solution $\boldsymbol{\sigma}_k$.
    Compute the error estimator $\eta = \eta(\boldsymbol{\sigma}_k, \mathcal{E}_k)$.
    Mark the minimal edge set $\mathcal{M}_k$ such that

$$\eta^2(\boldsymbol{\sigma}_k, \mathcal{M}_k) \geq \theta\, \eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k). \tag{6.1}$$

    If $\operatorname{osc}(f, \mathcal{T}_k) > \operatorname{osc}(f, \mathcal{T}_0)\mu^k$, enlarge $\mathcal{M}_k$ such that

$$\operatorname{osc}(f, \Omega_{\mathcal{M}_k}) \geq \tilde{\theta}\, \operatorname{osc}(f, \mathcal{T}_k). \tag{6.2}$$

    Refine each triangle $\tau \in \Omega_{\mathcal{M}_k}$ by the newest vertex bisection to get $\mathcal{T}_{k+1}$.
    $k = k + 1$.
  END WHILE
  $\mathcal{T}_N = \mathcal{T}_k$.
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:

$$e_k = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_k\|^2, E_k = \|\boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k\|^2, o_k = \operatorname{osc}^2(f, \mathcal{T}_k), \text{ and } \eta_k = \eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k).$$

Lemma 6.1. *One has the following inequalities*

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

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

$$e_k \leq C_1\, \eta_k + C_0\, o_k. \tag{6.5}$$

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

$$\beta\eta^2(\boldsymbol{\sigma}_{k+1}, \mathcal{E}_{k+1}) \leq \beta\eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_{k+1}) + E_k. \tag{6.6}$$

Let $\mathcal{N}_{k+1} = \mathcal{E}_{k+1}\backslash\mathcal{E}_k$ be the new edges in $\mathcal{T}_{k+1}$ and $\overline{\mathcal{M}}_k \subseteq \mathcal{E}_k$ be the refined edge in $\mathcal{T}_k$. It is obvious that $\mathcal{E}_k\backslash\overline{\mathcal{M}}_k = \mathcal{E}_{k+1}\backslash\mathcal{N}_{k+1}$. For an edge $E \in \mathcal{N}_{k+1}$, if it is an interior edge of some triangle $T \in \mathcal{T}_k$, then $J_E(\boldsymbol{\sigma}_k) = 0$ since $\boldsymbol{\sigma}_k$ is a polynomial in T. For other edges, it is at least half of some edge in $\overline{\mathcal{M}}_k$ and thus

$$\eta^2(\boldsymbol{\sigma}_k, \mathcal{N}_{k+1}) \leq \frac{1}{2}\,\eta^2(\boldsymbol{\sigma}_k, \overline{\mathcal{M}}_k). \tag{6.7}$$

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

$$\eta^2(\boldsymbol{\sigma}_k, \overline{\mathcal{M}}_k) \geq \eta^2(\boldsymbol{\sigma}_k, \mathcal{M}_k) \geq \theta\,\eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k). \tag{6.8}$$

Combining (6.7) and (6.8), we get

$$\begin{aligned}
\eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_{k+1}) &= \eta^2(\boldsymbol{\sigma}_k, \mathcal{N}_{k+1}) + \eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_{k+1}\backslash\mathcal{N}_{k+1}) \\
&\leq \frac{1}{2}\eta^2(\boldsymbol{\sigma}_k, \overline{\mathcal{M}}_k) + \eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k\backslash\overline{\mathcal{M}}_k) \\
&\leq -\frac{1}{2}\eta^2(\boldsymbol{\sigma}_k, \overline{\mathcal{M}}_k) + \eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k) \\
&\leq (1 - \frac{1}{2}\theta)\eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k).
\end{aligned}$$

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

Theorem 6.2. *When*

$$0 < \delta < \min\{\frac{\beta}{2C_1}\theta, 1\}, \tag{6.9}$$

there exists $\alpha \in (0, 1)$ and C_δ such that

$$(1 - \delta)e_{k+1} + \beta\eta_{k+1} \leq \alpha\big[(1 - \delta)e_k + \beta\eta_k\big] + C_\delta\,o_k. \tag{6.10}$$

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_k and use (6.4) to bound

$$\begin{aligned}
e_k &= \alpha(1 - \delta)e_k + [1 - \alpha(1 - \delta)]e_k \\
&\leq \alpha(1 - \delta)e_k + [1 - \alpha(1 - \delta)](C_1\eta_k + C_0\,o_k).
\end{aligned}$$

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 α 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 δ (6.9), we conclude $\alpha \in (0, 1)$. $\qquad\square$

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

$$(1 - \delta)\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_k\|^2 + \beta\eta^2(\boldsymbol{\sigma}_k, \mathcal{T}_k) \le C_\delta \gamma_\delta^k,$$

and thus the algorithm **AMFEM** *will terminate in finite steps.*

Proof. The proof is identical to that of Theorem 4.7 in [50] using (6.10). ☐

6.2. **Optimality of AMFEM.** Let $\mathcal{T}_0$ be an initial quasi-uniform triangulation with $\#\mathcal{T}_0 > 2$ and $\mathcal{P}_N$ be the set of all triangulations $\mathcal{T}$ which is refined from $\mathcal{T}_0$ and $\#\mathcal{T} \le N$. For a given triangulation $\mathcal{T}$, the solution of the mixed finite element approximation of Poisson equation will be denoted by $\boldsymbol{\sigma}_\mathcal{T}$. We define

$$\mathcal{A}^s = \{\boldsymbol{\sigma} \in \Sigma : \|\boldsymbol{\sigma}\|_{\mathcal{A}^s} < \infty, \text{ with } \|\boldsymbol{\sigma}\|_{\mathcal{A}^s} = \sup_{N \ge \#\mathcal{T}_0} \left(N^s \inf_{\mathcal{T} \in \mathcal{P}_N} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_\mathcal{T}\|\right)\}.$$

An adaptive finite element method realizes optimal convergence rates if whenever $\boldsymbol{\sigma} \in \mathcal{A}^s$, it produces approximation $\boldsymbol{\sigma}_N$ with respect to triangulations $\mathcal{T}_N$ elements such that $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_N\| \le C(\#\mathcal{T}_N)^{-s}$.

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

(1) $[\mathcal{T}_H, f_H] = $ **APPROX** $(f, \mathcal{T}_0, \varepsilon/2)$

(2) $[\boldsymbol{\sigma}_N, \mathcal{T}_N] = $ **AMFEM**$(\mathcal{T}_H, f_H, \varepsilon/2, \theta, 0, 1)$

The advantage of separating data error and discretization 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 $\mathcal{T}_H$ for **AMFEM**. In this case, we also list all ingredients needed for the optimality of adaptive procedure.

(1) Orthogonality:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{k+1}\|^2 = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_k\|^2 - \|\boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k\|^2$$

(2) Discrete upper bound:

$$\|\boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k\|^2 \le C_1\eta^2(\boldsymbol{\sigma}_k, \mathcal{F}_k) \text{ and } \#\mathcal{F}_k \le 3(\#\mathcal{T}_{k+1} - \#\mathcal{T}_k).$$

(3) Lower bound:

$$C_2\eta^2(\boldsymbol{\sigma}_k, \mathcal{E}_k) \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_k\|^2.$$

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

$$\|\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_N\| \le \varepsilon, \quad \text{and} \quad \#\mathcal{T}_N - \#\mathcal{T}_0 \le C\|\boldsymbol{\sigma}\|_{\mathcal{A}^s}^{1/s}\varepsilon^{-1/s}. \tag{6.11}$$

Proof. It is identical to the proof of Theorem 5.3 in [59] using three ingredients listed above. ☐

Theorem 6.5. *For any $f \in L^2(\Omega)$, a shape regular triangulation $\mathcal{T}_0$ and $\varepsilon > 0$. Let $\boldsymbol{\sigma} = \mathcal{L}^{-1}f$ and $[\boldsymbol{\sigma}_N, \mathcal{T}_N] = $ **AMFEM** $(\mathcal{T}_H, f_H, \varepsilon/2, 0, 1)$ where $[\mathcal{T}_H, f_H] = $ **APPROX** $(f, \mathcal{T}_0, \varepsilon/2)$. If $\boldsymbol{\sigma} \in \mathcal{A}^s$ and $f \in \mathcal{A}_o^s$, then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_N\| \le C\big(\|\boldsymbol{\sigma}\|_{\mathcal{A}^s} + \|f\|_{\mathcal{A}_o^s}\big)(\#\mathcal{T}_N - \#\mathcal{T}_0)^{-s}.$$

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

$$\|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\| \leq \varepsilon/2, \quad \text{and} \quad \#\mathcal{T}_H - \#\mathcal{T}_0 \leq C\|f\|_{\mathcal{A}_o^s}^{1/s}\varepsilon^{-1/s}. \tag{6.12}$$

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

$$\|\tilde{\boldsymbol{\sigma}}\|_{\mathcal{A}^s} \leq \|\boldsymbol{\sigma}\|_{\mathcal{A}^s} + \|f\|_{\mathcal{A}_o^s}$$

We then apply Theorem 6.4 to $\tilde{\boldsymbol{\sigma}}$ to obtain

$$\|\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_N\| \leq \varepsilon/2 \quad \text{and} \quad \#\mathcal{T}_N - \#\mathcal{T}_H \leq C\|\tilde{\boldsymbol{\sigma}}\|_{\mathcal{A}^s}^{1/s}\varepsilon^{-1/s}. \tag{6.13}$$

Combining (6.12) and (6.13) we get

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_N\| \leq \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\| + \|\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_N\| \leq \varepsilon$$

and

$$\varepsilon \leq C(\#\mathcal{T}_N - \#\mathcal{T}_0)^{-s}\big(\|\boldsymbol{\sigma}\|_{\mathcal{A}^s} + \|f\|_{\mathcal{A}_o^s}\big).$$

The desired result then follows.   □

7. CONCLUSION AND FUTURE WORK

In this paper, we have designed and analyzed convergent adaptive mixed finite element methods with optimal complexity for arbitrary order Raviart-Thomas and Brezzi-Douglas-Marini elements. Although the results are presented in two dimensions, most of them are dimensional independent. For example, the discrete stability result, Theorem 4.1, holds in arbitrary dimensions without any modification of the proof.

The proof for the upper bound of the error estimator (Theorem 5.3 and 5.7), however, cannot be generalized to three dimensions in a straightforward way. In the proof, we use a special fact that in two dimensions, $H(\text{curl})$ is as smooth as H^1 since in two dimensions curl operator is just a rotation of gradient operator. To overcome this difficulty, we need to use a regular decomposition instead of Helmholtz decomposition. Note that discrete regular decomposition for corresponding finite element spaces is developed recently by Hiptmair and Xu [42]. We could use these techniques to prove the convergence and optimality of adaptive mixed finite element methods in three and higher dimensions.

REFERENCES

[1] A. Alonso. Error estimators for a mixed method. *Numerische Mathematik*, 74(4):385–395, 1996.

[2] D. N. Arnold. Differential complexes and numerical stability. *Plenary address delivered at ICM 2002 International Congress of Mathematicians*, 2004.

[3] D. N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: Implementation, postporcessing and error estimates. *RAIRO Model Math. Anal. Numer.*, 19:7–32, 1985.

[4] D. N. Arnold, R. S. Falk, and R. Winther. Multigrid in $H(div)$ and $H(curl)$. *Numerische Mathematik*, 85:197–218, 2000.

[5] D. N. Arnold, R. S. Falk, and R. Winther. Differential complexes and stability of finite element methods. I. the de Rham complex. *preprint*, 2005.

[6] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numerica*, pages 1–155, 2006.

[7] D. N. Arnold, L. R. Scott, and M. Vogelius. Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(2):169–192 (1989), 1988.

[8] I. Babuška and M. Vogelius. Feeback and adaptive finite element solution of one-dimensional boundary value problems. *Numerische Mathematik*, 44:75–102, 1984.

[9] C. Bacuta, J. H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems in besov spaes. *Mathematics of Computation*, 72(244):1577–1595, 2002.

[10] C. Bacuta, J. H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems with smooth data on polygonal domains. *Numerische Mathematik*, 11(2):75–94, 2003.

[11] R. E. Bank, A. H. Sherman, and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In *Scientific Computing*, pages 3–17. IMACS/North-Holland Publishing Company, Amsterdam, 1983.

[12] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97(2):219–268, 2004.

[13] P. Binev, W. Dahmen, R. DeVore, and P. Petrushev. Approximation classes for adaptive methods. *Serdica Math. J*, 28:391–416, 2002.

[14] P. Binev and R. DeVore. Fast computation in adaptive tree approximation. *Numerische Mathematik*, 97:193–217, 2004.

[15] P. Bochev and M. Gunzburger. On least-sequares finite element methods for the poisson equation and their connection to the Dirichlet and Kelvin principles. *SIAM Journal on Numerical Analysis*, 43(1):340–362, 2005.

[16] A. Bossavit. Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism. *Science, Measurement and Technology, IEE Proceedings*, 135(8):493–500, Nov 1988.

[17] D. Braess and R. Verfürth. Multigrid methods for nonconforming finite element methods. *SIAM Journal on Numerical Analysis*, 27:979–986, 1990.

[18] J. H. Bramble, J. E. Pasciak, and J. Xu. The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems. *Mathematics of Computation*, 51:389–414, 1988.

[19] J. H. Brandts. *Superconvergence phenomena in finite element methods*. PhD thesis, Utrecht University, 1994.

[20] S. C. Brenner. A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method. *SIAM Journal on Numerical Analysis*, 29:647–678, 1992.

[21] S. C. Brenner. Two-level additive Schwarz preconditioners for nonconforming finite element methods. *Mathematics of Computation*, 65:897–921, 1996.

[22] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2002.

[23] F. Brezzi, J. Douglas, and L. D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numerische Mathematik*, 47(2):217–235, 1985.

[24] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, 1991.

[25] C. Carstensen. A posteriori error estimate for the mixed finite element method. *Mathematics of Computation*, 66:465–476, 1997.

[26] C. Carstensen and R. H. W. Hoppe. Error reduction and convergence for an adaptive mixed finite element method. *Math. Comp.*, 75(255):1033–1042 (electronic), 2006.

[27] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *Preprint 9, University of Augsburg*, 2007.

[28] L. Chen. Short implementation of bisection in MATLAB. *report*, 2006.

[29] L. Chen and J. Xu. Topics on adaptive finite element methods. In T. Tang and J. Xu, editors, *Adaptive Computations: Theory and Algorithms*, pages 1–31. Science Press, Beijing, 2007.

[30] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 4 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.

[31] B. Cockburn and J. Gopalakrishnan. A characterization of hybridized mixed methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 42(1):283–301, 2004.

[32] B. Cockburn and J. Gopalakrishnan. Error analysis of variable degree mixed methods for elliptic problems via hybridization. *Mathematics of Computation*, 74(252):1653–1677, 2005.

[33] B. Cockburn and J. Gopalakrishnan. New hybridization techniques. *GAMM-Mitt.*, 28(2):154–182, 2005.

[34] S. Dahlke. Besov regularity for elliptic boundary value problems on polygonal domains. *Appl. Math. Lett.*, 12:31–36, 1999.

[35] S. Dahlke and R. A. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1&2):1–16, 1997.

[36] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33:1106–1124, 1996.

[37] R. G. Durán and M. A. Muschietti. An explicit right inverse of the divergence operator which is continuous in weighted norms. *Studia Math.*, 148(3):207–219, 2001.

[38] G. J. Fix, M. D. Gunzburger, and R. A. Nicolaides. On mixed finite element methods for first order elliptic systems. *Numerische Mathematik*, 37(1):29–48, 1981.

[39] G. N. Gatica and M. Maischak. A posteriori error estimates for the mixed finite element method with lagrange multipliers. *Numer. Methods Partial Differential Equations*, 21(3):421 – 450, 2004.

[40] J. Gopalakrishnan. A Schwarz preconditioner for a hybridized mixed method. *Computational Methods In Applied Mathematics*, 3(1):116—134, 2003.

[41] R. Hiptmair. Canonical construction of finite elements. *Mathematics of Computation*, 68:1325–1346, 1999.

[42] R. Hiptmair and J. Xu. Nodal auxiliary space preconditioning in H(curl) and H(div) spaces. Research report no. 2006-09, ETH, Zurich, Switzerland, 2006.

[43] R. H. W. Hoppe and B. Wohlmuth. Adaptive multilevel techniques for mixed finite element discretizations of elliptic boundary value problems. *SIAM Journal on Numerical Analysis*, 34(4):1658–1681, aug 1997.

[44] M. G. Larson and A. Maqvist. A posteriori error estimates for mixed finite element approximations of elliptic problems. *Preprint*, 2005.

[45] C. Lovadina and R. Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Mathemathics of Computation*, Primary 65N30, 2006.

[46] W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software (TOMS) archive*, 15(4):326 – 347, 1989.

[47] W. F. Mitchell. Optimal multilevel iterative methods for adaptive grids. *SIAM Journal on Scientific and Statistical Computing*, 13:146–167, 1992.

[48] P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM Journal on Numerical Analysis*, 38(2):466–488, 2000.

[49] P. Morin, R. H. Nochetto, and K. G. Siebert. Convergence of adaptive finite element methods. *SIAM Review*, 44(4):631–658, 2002.

[50] P. Morin, R. H. Nochetto, and K. G. Siebert. Local problems on stars: A posteriori error estimators, convergence, and performance. *Mathematics of Computation*, 72:1067–1097, 2003.

[51] P. Morin, K. G. Siebert, and A. Veeser. A basic convergence result for conforming adaptive finite elements. *Preprint, University of Augsburg*, 2007.

[52] P. Oswald. Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations. *Applied Numerical Mathematics*, 23(1):139–158, 1997.

[53] P. A. Raviart and J. Thomas. A mixed finite element method fo 2-nd order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical aspects of the Finite Elements Method*, Lectures Notes in Math. 606, pages 292–315. Springer, Berlin, 1977.

[54] M. C. Rivara. Design and data structure for fully adaptive, multigrid finite element software. *ACM Trans. Math. Soft.*, 10:242–264, 1984.

[55] M. C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM Journal on Numerical Analysis*, 21:604–613, 1984.

[56] T. Rusten, P. S. Vassilevski, and R. Winther. Interior penalty preconditioners for mixed finite element approximations of elliptic problems. *Mathemathics of Computation*, 65:447–466, 1996.

[57] R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483–493, 1990.

[58] E. G. Sewell. Automatic generation of triangulations for piecewise polynomial approximation. In *Ph. D. dissertation*. Purdue Univ., West Lafayette, Ind., 1972.

[59] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

[60] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Mathemathics of Computation*, 77:227–241, 2008.

[61] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh refinement tecniques*. B. G. Teubner, 1996.

[62] B. I. Wohlmuth and R. H. W. Hoppe. A comparison of a posteriori error estimators for mixed finite element discretizations by raviart-thomas elements. *Mathematics of Computation*, 82:253–279, 1999.

E-mail address: chenlong@math.uci.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT IRVINE, IRVINE, CA 92697

E-mail address: mholst@math.ucsd.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT SAN DIEGO, LA JOLLA, CA 92093

E-mail address: xu@math.psu.edu

THE SCHOOL OF MATHEMATICAL SCIENCE, PEKING UNIVERSITY, AND DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16801