arXiv:1601.00132v1 [math.NA] 2 Jan 2016

A UNIFIED ANALYSIS OF QUASI-OPTIMAL CONVERGENCE FOR ADAPTIVE MIXED FINITE ELEMENT METHODS

JUN HU[1] AND GUOZHU YU[2,*]

[1]*LMAM and School of Mathematical Sciences, Peking University, Beijing 100871,China*
[2]*School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China*

Abstract. In this paper, we present a unified analysis of both convergence and optimality of adaptive mixed finite element methods for a class of problems when the finite element spaces and corresponding a posteriori error estimates under consideration satisfy five hypotheses. We prove that these five conditions are sufficient for convergence and optimality of the adaptive algorithms under consideration. The main ingredient for the analysis is a new method to analyze both discrete reliability and quasi-orthogonality. This new method arises from an appropriate and natural choice of the norms for both the discrete displacement and stress spaces, namely, a mesh-dependent discrete H^1 norm for the former and a L^2 norm for the latter, and a newly defined projection operator from the discrete stress space on the coarser mesh onto the discrete divergence free space on the finer mesh. As applications, we prove these five hypotheses for the Raviart–Thomas and Brezzi–Douglas–Marini elements of the Poisson and Stokes problems in both 2D and 3D.

1. Introduction

This paper is devoted to convergence and optimality of adaptive mixed finite element methods (AMFEMs) for the problem of the following form: Given $f \in L^2(\Omega)$, find $(\sigma, u) \in \Sigma \times U$ such that

$$(\mathcal{A}\sigma, \tau)_{L^2(\Omega)} - (\operatorname{div}\tau, u)_{L^2(\Omega)} = 0, \text{ for any } \tau \in \Sigma,$$
$$(\operatorname{div}\sigma, v)_{L^2(\Omega)} - (f, v)_{L^2(\Omega)} = 0, \text{ for any } v \in U. \tag{1.1}$$

In the paper, we refer to Σ as the stress space, and U as the displacement space. Here, Ω is a simply connected bounded domain in $\mathbb{R}^d (d = 2, 3)$ with the boundary $\partial\Omega$, and Σ, U are Sobolev spaces defined as

$$\Sigma := H(\operatorname{div}, \Omega; \mathbb{R}^{d\times n}), \quad U := L^2(\Omega; \mathbb{R}^n)$$

with n some positive integer. Furthermore, we assume $\mathcal{A}$ is a linear, bounded and semi-definite operator, satisfying

$$0 \le (\mathcal{A}\tau, \tau)_{L^2(\Omega)} \le C\|\tau\|^2_{L^2(\Omega)} \text{ and } \|\tau\|^2_{L^2(\Omega)} \le C((\mathcal{A}\tau, \tau)_{L^2(\Omega)} + \|\operatorname{div}\tau\|^2_{H^{-1}(\Omega)}) \tag{1.2}$$

*: Corresponding author.
The research of the first author was supported by NSFC projection 11271035, 91430213 and 11421101.
The research of the second author was supported by NSFC projection 11401492.
Email addresses: hujun@math.pku.edu.cn (J. Hu), yuguozhumail@163.com (G. Yu).

for any $\tau \in \Sigma$, with the $H^{-1}(\Omega)$ norm defined as

$$\|\psi\|_{H^{-1}(\Omega)} := \sup_{v\in H_0^1(\Omega)} \frac{(\psi, v)}{\|v\|_{H^1(\Omega)}}.$$

For convenience, we will also use $\|\tau\|_{\mathcal{A}}$ to denote $(\mathcal{A}\tau,\tau)^{1/2}_{L^2(\Omega)}$ when there is no confusion.

Many problems can be attributed to the form of (1.1). For example, when

$$\mathcal{A}\tau := \tau, \quad \Sigma \times U := H(\mathrm{div}, \Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}),$$

the problem (1.1) is essentially the mixed formulation of the Poisson problem; when

$$\mathcal{A}\tau := \tau - \frac{1}{d}(\mathrm{tr}\,\tau) I_{d\times d}, \quad \Sigma \times U := \left\{\tau \in H(\mathrm{div}, \Omega; \mathbb{R}^{d\times d}) | \int_\Omega \mathrm{tr}\,\tau dx = 0\right\} \times L^2(\Omega; \mathbb{R}^d)$$

with the $d \times d$ identity matrix $I_{d\times d}$, then the problem (1.1) becomes the pseudostress-velocity formulation of the stationary Stokes problem, see for instance, [4, 11, 12, 19] and the references therein. Here and throughout this paper, the trace operator tr is defined as

$$\mathrm{tr}\,\tau = \sum_{i=1}^d \tau_{ii} \text{ for any matrix } \tau \in \mathbb{R}^{d\times d}.$$

For the problem (1.1), the theory of the reliable and efficient a posteriori error analysis has been in some sense relatively mature. We refer the interested readers to [9, 2, 8, 13, 42, 41, 32, 1, 31, 30] and the references therein for a posteriori error estimates of the mixed finite element methods of the Poisson problem and [19] for a posteriori error estimates of the mixed finite element methods of the Stokes problem within the pseudostress-velocity formulation. We also mention the references [26, 14, 40, 15, 25] for the other related works.

As for the convergence and optimality analysis, there have been several results for adaptive conforming and nonconforming finite element methods [5, 23, 34, 35, 7, 33, 28, 39, 21, 36, 27, 16]; while for AMFEMs, research efforts are made mainly on the Poisson problem. Carstensen and Hoppe [18] established the first error reduction and convergence of the adaptive lowest-order Raviart–Thomas element method, and similar results can be found in [6, 20]. Later, in [22, 24] convergence and optimality were analyzed for the Raviart–Thomas and Brezzi–Douglas–Marini elements of any order. By using the discrete Helmholtz decomposition, Huang and Xu [29] extended the above results to the 3D case. For the mixed finite elements of the Stokes problem within the pseudostress-velocity formulation, Carstensen et. al [17] proved convergence and optimality of the adaptive lowest-order Raviart–Thomas element. The main ingredients therein are some novel equivalence between the lowest-order Raviart–Thomas and Crouzeix–Raviart elements, and a particular Helmholtz decomposition of deviatoric tensors for the 2D case. However, the analysis can neither be generalized to the Raviart–Thomas and Brezzi–Douglas–Marini elements of any order, nor to the 3D case. We refer interested readers to [16] for a comprehensive review of the state of art of this field and also a wonderful simultaneous axiomatic analysis of both convergence and optimality of adaptive finite element methods of several classes of linear and nonlinear problems.

This paper aims at a unified convergence and optimality analysis of AMFEMs of the problem (1.1) in both two and three dimensional cases. The main result states that if the mixed finite element methods and associated a posteriori error estimates satisfy five hypotheses, see more details in next section, the corresponding adaptive algorithms converge with optimal rates in the nonlinear approximate sense. The unified analysis is based on a new method to establish both discrete reliability and quasi-orthogonality, which are two main and indispensable ingredients for the convergence and optimality analysis of adaptive finite element methods. In fact, in contrary to [18, 6, 22, 20, 24, 29, 17], the discrete displacement space is endowed with a mesh-dependent discrete H^1 norm, which defines one component of the new method. Hence the L^2 norm becomes a natural norm for the discrete stress space. The other component of the new method is to introduce a projection operator from the discrete stress space on the coarser mesh onto the discrete divergence free space on the finer mesh. As applications, the Raviart–Thomas and Brezzi–Douglas–Marini elements of any order of both the Poisson problem and the Stokes problem within the pseudostress-velocity formulation in 2D and 3D are proved to satisfy these five hypotheses. Therefore the corresponding adaptive schemes admit optimal convergence. As a result, it extends the optimal convergence result for the first order Raviart–Thomas element of the Stokes problem in 2D from [17] to the more general case.

Throughout this paper, the notation $a \lesssim b$ represents that there exists a generic positive constant C, which is independent of the mesh parameter h and may not be the same at different occurrences, such that $a \leq Cb$. The symbol $a \approx b$ means $a \lesssim b \lesssim a$.

Let v, $\beta = (\beta_1, \cdots, \beta_d)^T$ and $\tau = (\tau_{ij})_{d\times d}$ be scalar, vector and tensor functions of two or three variables respectively, and let $\tau_i = (\tau_{i1}, \cdots, \tau_{id})^T$ denote the ith row for τ with $i = 1, \cdots, d$. We define the grad, div, curl and rot operators by

$$
\begin{gathered}
\operatorname{grad} v := \left(\frac{\partial v}{\partial x_1}, \cdots, \frac{\partial v}{\partial x_d}\right)^T, \quad \operatorname{grad} \beta := (\operatorname{grad} \beta_1, \cdots, \operatorname{grad} \beta_d)^T, \\
\operatorname{div} \beta := \frac{\partial \beta_1}{\partial x_1} + \cdots + \frac{\partial \beta_d}{\partial x_d}, \quad \operatorname{div} \tau := (\operatorname{div} \tau_1, \cdots, \operatorname{div} \tau_d)^T, \\
\operatorname{curl} v := \left(\frac{\partial v}{\partial x_2}, -\frac{\partial v}{\partial x_1}\right)^T, \quad \operatorname{curl} \beta := (\operatorname{curl} \beta_1, \operatorname{curl} \beta_2)^T, \quad d = 2, \\
\operatorname{curl} \beta := \operatorname{grad} \times \beta, \quad \operatorname{curl} \tau := (\operatorname{curl} \tau_1, \operatorname{curl} \tau_2, \operatorname{curl} \tau_3)^T, \quad d = 3, \\
\operatorname{rot} \beta := \frac{\partial \beta_1}{\partial x_2} - \frac{\partial \beta_2}{\partial x_1}, \quad \operatorname{rot} \tau := (\operatorname{rot} \tau_1, \operatorname{rot} \tau_2)^T, \quad d = 2.
\end{gathered}
$$

Moreover, for the tensor function τ, we define its tangential component by

$$
\tau \cdot t := (\tau_1 \cdot t, \tau_2 \cdot t)^T \text{ for } d = 2, \quad \tau \times \nu := (\tau_1 \times \nu, \tau_2 \times \nu, \tau_3 \times \nu)^T \text{ for } d = 3.
$$

For a given Lebesgue measurable set $G \subset \mathbb{R}^d$, we use $L^2(G; \mathbb{R})$ or $L^2(G; \mathbb{R}^{d\times n})$ to denote the Hilbert space of square integrable functions or matrix-value fields, respectively, with inner product $(\cdot, \cdot)_{L^2(G)}$. Here and thereafter we will omit $\mathbb{R}$ or $\mathbb{R}^{d\times n}$ for simplicity when there is no risk of confusion.

We also define the following spaces

$$
\begin{aligned}
H^1(G) &:= \{v \in L^2(G) | \operatorname{grad} v \in L^2(G)\}, \\
H(\operatorname{div}, G) &:= \{v \in L^2(G) | \operatorname{div} v \in L^2(G)\}, \\
H(\operatorname{curl}, G) &:= \{v \in L^2(G) | \operatorname{curl} v \in L^2(G)\},
\end{aligned}
$$

equipped with norms

$$\|v\|_{H^1(G)} := (\|v\|^2_{L^2(G)} + \|\operatorname{grad} v\|^2_{L^2(G)})^{1/2}, \quad \text{for all } v \in H^1(G),$$

$$\|v\|_{H(\operatorname{div},G)} := (\|v\|^2_{L^2(G)} + \|\operatorname{div} v\|^2_{L^2(G)})^{1/2}, \quad \text{for all } v \in H(\operatorname{div}, G),$$

$$\|v\|_{H(\operatorname{curl},G)} := (\|v\|^2_{L^2(G)} + \|\operatorname{curl} v\|^2_{L^2(G)})^{1/2}, \quad \text{for all } v \in H(\operatorname{curl}, G),$$

respectively, where $\|\cdot\|_{L^2(G)} := (\cdot,\cdot)^{1/2}_{L^2(G)}$ denotes the norm of space $L^2(G)$. Especially, let $H^1_0(G) := \{v \in H^1(G), v|_{\partial G} = 0\}$.

The rest of the paper is organized as follows. In Section 2, we present notation and five hypotheses. In Section 3, we show discrete reliability and quasi-orthogonality under these five hypotheses. In Section 4, we prove convergence and optimality of the adaptive algorithms while in Section 5 we check these five hypotheses for two examples.

2. Notation and Hypothesis

Let $\mathcal{T}_h$ be some shape-regular triangulation of Ω and $\mathcal{E}_h$ the set of all edges or faces in $\mathcal{T}_h$. We indicate by $h_K := |K|^{1/d}$ and $h_E := |E|^{1/(d-1)}$ the size for each $K \in \mathcal{T}_h$ and each $E \in \mathcal{E}_h$, respectively. Note that all geometric entities are closed sets. Given $K \in \mathcal{T}_h$ and $E \in \mathcal{E}_h$, define the element-patch and the edge/face-patch by

$$\Omega_K := \bigcup_{K' \cap K \neq \emptyset} K', \quad \Omega_E := \bigcup_{E \subset K} K,$$

respectively. Given any interior edge/face $E \in \mathcal{E}_h$, let ν_E be a unit normal vector, and $[\cdot]|_E := \cdot|_{K^+} - \cdot|_{K^-}$ be the jump across the edge/face $E = K^+ \cap K^-$ shared by the two elements $K^+, K^- \in \mathcal{T}_h$. While for the boundary edge/face E, ν_E denotes the unit outer vector normal to $\partial\Omega$, and the jump $[\cdot]|_E := \cdot|_{K^+}$ for the unique element K^+ with $E \subset K^+$. When $d = 2$ let t_E be the unit tangential vector of E.

Let $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$, and $\mathcal{R} := \mathcal{T}_H \backslash \mathcal{T}_h = \{K \in \mathcal{T}_H | K \notin \mathcal{T}_h\}$ be the set of refined elements from $\mathcal{T}_H$ to $\mathcal{T}_h$, $\tilde{\mathcal{R}} := \{K \in \mathcal{T}_H | K \cap K' \neq \emptyset \text{ for some } K' \in \mathcal{R}\}$. Also, let U_h and U_H, Σ_h and Σ_H, $H_h(\operatorname{curl}, \Omega)$ and $H_H(\operatorname{curl}, \Omega)$ be finite element subspaces of U, Σ and $H(\operatorname{curl}, \Omega)$ defined on $\mathcal{T}_h$ and $\mathcal{T}_H$, respectively.

The mixed finite element method is to solve (1.1) in the pair of the finite dimensional spaces $\Sigma_h \times U_h \subset \Sigma \times U$. The corresponding discrete problem reads: Find $(\sigma_h, u_h) \in \Sigma_h \times U_h$ such that

$$\begin{aligned} &(\mathcal{A}\,\sigma_h, \tau_h)_{L^2(\Omega)} - (\operatorname{div} \tau_h, u_h)_{L^2(\Omega)} = 0, \text{ for any } \tau_h \in \Sigma_h, \\ &(\operatorname{div} \sigma_h, v_h)_{L^2(\Omega)} - (f, v_h)_{L^2(\Omega)} = 0, \text{ for any } v_h \in U_h. \end{aligned} \tag{2.1}$$

Let $\mathcal{Q}_h$ be the L^2-projection operator from U onto the space U_h. The edge or face error estimator with respect to a given subset $\mathcal{M}_h \subseteq \mathcal{T}_h$ is defined by, [19, 29],

$$\eta^2(\sigma_h, \mathcal{M}_h) := \begin{cases} \sum_{K \in \mathcal{M}_h} \left(\|h_K \operatorname{rot}(\mathcal{A}\,\sigma_h)\|^2_{L^2(K)} + \sum_{E \in K \cap \mathcal{E}_h} \|h_E^{1/2} [\mathcal{A}\,\sigma_h \cdot t_E]\|^2_{L^2(E)} \right) & d = 2, \\ \sum_{K \in \mathcal{M}_h} \left(\|h_K \operatorname{curl}(\mathcal{A}\,\sigma_h)\|^2_{L^2(K)} + \sum_{E \in K \cap \mathcal{E}_h} \|h_E^{1/2} [\mathcal{A}\,\sigma_h \times \nu_E]\|^2_{L^2(E)} \right) & d = 3. \end{cases}$$

And given $f \in L^2(\Omega)$, define the data oscillation by

$$\operatorname{osc}^2(f, \mathcal{M}_h) := \sum_{K \in \mathcal{M}_h} \|h_K(f - f_h)\|^2_{L^2(K)} \text{ with } f_h = \mathcal{Q}_h f.$$

It follows that

$$\operatorname{osc}^2(f_h, \mathcal{T}_H) = \sum_{K \in \mathcal{T}_H \setminus \mathcal{T}_h} \|h_K(f_h - f_H)\|^2_{L^2(K)}.$$

For each $K \in \mathcal{T}_H$, since $\|h_K(f_h - f_H)\|_{L^2(K)} = \|h_K \mathcal{Q}_h(f - f_H)\|_{L^2(K)} \leq \|h_K(f - f_H)\|_{L^2(K)}$,

$$\operatorname{osc}^2(f_h, \mathcal{T}_H) \leq \sum_{K \in \mathcal{T}_H \setminus \mathcal{T}_h} \|h_K(f - f_H)\|^2_{L^2(K)} = \operatorname{osc}^2(f, \mathcal{T}_H \setminus \mathcal{T}_h).$$

Given $f \in L^2(\Omega)$, define an affine space $Z_h(f)$ by

$$Z_h(f) := \{\tau_h \in \Sigma_h, \operatorname{div} \tau_h = f_h\}. \tag{2.2}$$

In particular, $Z_h(0)$ is the kernel space of the discrete divergence operator, which is also called the discrete divergence free space. Let (σ_h, u_h) be the solution of (2.1), we then have the following key property

$$(\mathcal{A}\,\sigma_h, \tau_h)_{L^2(\Omega)} = 0, \text{ for any } \tau_h \in Z_h(0). \tag{2.3}$$

We follow [3, Lemma 2.1] to endow the space U_h with the following discrete H^1 norm: for a given set G consisting of elements $K \in \mathcal{T}_h$,

$$\|v_h\|^2_{1,h,G} := \sum_{K \in G \cap \mathcal{T}_h} \|\nabla v_h\|^2_{L^2(K)} + \sum_{E \in G \cap \mathcal{E}_h} h_E^{-1} \|[v_h]\|^2_{L^2(E)}. \tag{2.4}$$

When $G = \Omega$, the subscript is omitted. Hence the naturally matched norm for the space Σ_h is the the L^2 norm.

Next, we propose five hypotheses, which are sufficient for convergence and optimality of AMFEMs.

Hypothesis 1. *The discrete spaces Σ_h and U_h satisfy the following inclusion properties*

$$\Sigma_H \subset \Sigma_h \text{ and } \operatorname{div} \Sigma_h \subset U_h.$$

Hypothesis 2. *The pair of spaces (Σ_h, U_h) satisfies the discrete inf-sup condition*

$$\|v_h\|_{1,h} \lesssim \sup_{0 \neq \tau_h \in \Sigma_h} \frac{(\operatorname{div} \tau_h, v_h)_{L^2(\Omega)}}{\|\tau_h\|_{L^2(\Omega)}}, \text{ for any } v_h \in U_h,$$

which implies the following equivalent inf-sup condition, see Brezzi and Fortin [10] *for more details,*

$$\|\tau_h\|_{L^2(\Omega)/Z_h(0)} \lesssim \sup_{0 \neq v_h \in U_h} \frac{(\operatorname{div} \tau_h, v_h)_{L^2(\Omega)}}{\|v_h\|_{1,h}}, \text{ for any } \tau_h \in \Sigma_h. \tag{2.5}$$

Hypothesis 3. *Given $v_h \in U_h$, there exists an operator $\mathcal{S}_H : U_h \to U_H$ such that*

$$\mathcal{S}_H v_h|_K = v_h|_K, \text{ for any } K \in \mathcal{T}_H \cap \mathcal{T}_h, \tag{2.6}$$

and

$$\|v_h - \mathcal{S}_H v_h\|_{L^2(K)} \lesssim h_K \|v_h\|_{1,h,D_K}, \text{ for any } K \in \mathcal{T}_H \setminus \mathcal{T}_h, \tag{2.7}$$

where $D_K := \bigcup_{K' \in \mathcal{T}_h, K' \cap K \neq \emptyset} K'$.

Hypothesis 4. *Given* $\xi_h \in \Sigma_h$ *with* $\operatorname{div}\xi_h = 0$, *there exist* $\varphi_h \in H_h(\mathrm{curl},\Omega)$ *and an operator* $\Pi_H : H_h(\mathrm{curl},\Omega) \to H_H(\mathrm{curl},\Omega)$ *such that*

$$\xi_h = \mathrm{curl}\,\varphi_h \text{ and } \mathrm{curl}\,\Pi_H\varphi_h \in \Sigma_H. \tag{2.8}$$

Moreover, there exist $\psi \in H^1(\Omega)$ *and* $\phi \in L^2(\Omega)$ *such that*

$$\mathrm{curl}(\varphi_h - \Pi_H\varphi_h) = \mathrm{curl}\,\psi + \phi \tag{2.9}$$

with $(\mathcal{A}\sigma_H, \phi)_{L^2(\Omega)} = 0$ *(*(σ_H, u_H) *is the solution to the discrete problem* (2.1) *over* $\mathcal{T}_H$*) and*

$$\begin{cases} \psi|_K = 0, \quad \text{for any } K \in \mathcal{T}_H\backslash\tilde{\mathcal{R}}, \\ \left(\sum_{K\in\mathcal{T}_H} \|h_K^{-1}\psi\|^2_{L^2(K)}\right)^{1/2} \lesssim \|\,\mathrm{curl}\,\varphi_h\|_{L^2(\Omega)}, \\ \left(\sum_{E\in\mathcal{E}_H} \|h_E^{-1/2}\psi\|^2_{L^2(E)}\right)^{1/2} \lesssim \|\,\mathrm{curl}\,\varphi_h\|_{L^2(\Omega)}. \end{cases} \tag{2.10}$$

Besides these hypotheses on the finite element subspaces, the a posteriori error estimator with reliability and efficiency is necessary in an adaptive algorithm, which will be described in the following Hypothesis 5.

Hypothesis 5. *Let* (σ, u) *be the solution of* (1.1) *and* (σ_h, u_h) *be the solution of* (2.1) *over a triangulation* $\mathcal{T}_h$, *there exist constants* C_{Rel} *and* C_{Eff} *depending on the shape regularity of* $\mathcal{T}_h$ *such that*

$$\|\sigma - \sigma_h\|^2_{\mathcal{A}} \le C_{Rel}(\eta^2(\sigma_h, \mathcal{T}_h) + \mathrm{osc}^2(f, \mathcal{T}_h)), \quad \textit{(Reliability)} \tag{2.11}$$

$$C_{Eff}\eta^2(\sigma_h, \mathcal{T}_h) \le \|\sigma - \sigma_h\|^2_{\mathcal{A}}. \quad \textit{(Efficiency)} \tag{2.12}$$

We will in Section 5 show that the Raviart–Thomas and the Brezzi–Douglas–Marini elements for the Poisson and Stokes problems satisfy Hypotheses 1–5 in both two and three dimensions. We assume in the next two sections that these five hypotheses hold.

3. Discrete reliability and quasi-orthogonality

In this section we analyze discrete reliability of the estimator η, and also show quasi-orthogonality under the previous hypotheses. Compared to the analysis of both discrete reliability and quasi-orthogonality in literature, see for instance, [6, 22, 20, 24, 29, 17], the novelty of the analysis here is to equip the discrete displacement space U_h with the discrete H^1 norm defined in (2.4). Then it is natural to endow the discrete stress space Σ_h with the L^2 norm. Moreover, it allows us to make use of the equivalent form of the inf-sup condition (2.5).

Theorem 3.1. *Given* $f \in L^2(\Omega)$, *let* (σ_h, u_h) *and* (σ_H, u_H) *be the solutions to the discrete problem* (2.1) *over the nested triangulations* $\mathcal{T}_h$ *and* $\mathcal{T}_H$ *respectively. Then there exists a constant* C_{Drel} *such that*

$$\|\sigma_h - \sigma_H\|^2_{\mathcal{A}} \le C_{Drel}\left(\eta^2(\sigma_H, \tilde{\mathcal{R}}) + \mathrm{osc}^2(f, \mathcal{T}_H\backslash\mathcal{T}_h)\right). \tag{3.1}$$

Proof. The main idea is to evoke Hypothesis 2 and Hypothesis 4 to control the discrete divergence free part and its orthogonal complementary, separately. To this end, let ξ_h be the $\mathcal{A}$ projection of σ_H onto the space $Z_h(0)$ with

$$(\mathcal{A}\,\xi_h, \tau_h)_{L^2(\Omega)} = (\mathcal{A}\,\sigma_H, \tau_h)_{L^2(\Omega)}, \text{ for any } \tau_h \in Z_h(0).$$

By (2.3), the error $\|\sigma_h - \sigma_H\|_{\mathcal{A}}^2$ admits the following decomposition:

$$\|\sigma_h - \sigma_H\|_{\mathcal{A}}^2 = \|\sigma_h - \sigma_H + \xi_h - \xi_h\|_{\mathcal{A}}^2 \lesssim \|\sigma_h - \sigma_H + \xi_h\|_{\mathcal{A}}^2 + \|\xi_h\|_{\mathcal{A}}^2. \tag{3.2}$$

Since

$$(\mathcal{A}(\sigma_h - \sigma_H + \xi_h), \tau_h)_{L^2(\Omega)} = (\mathcal{A}\,\sigma_h, \tau_h)_{L^2(\Omega)} - (\mathcal{A}(\sigma_H - \xi_h), \tau_h)_{L^2(\Omega)} = 0,$$

for any $\tau_h \in Z_h(0)$, it follows from (1.2) and Hypothesis 2 that

$$\begin{aligned}
\|\sigma_h - \sigma_H + \xi_h\|_{\mathcal{A}} &\lesssim \|\sigma_h - \sigma_H + \xi_h\|_{L^2(\Omega)/Z_h(0)} \\
&\lesssim \sup_{v_h \in U_h} \frac{(\operatorname{div}(\sigma_h - \sigma_H + \xi_h), v_h)_{L^2(\Omega)}}{\|v_h\|_{1,h}} \\
&= \sup_{v_h \in U_h} \frac{(f_h - f_H, v_h)_{L^2(\Omega)}}{\|v_h\|_{1,h}}.
\end{aligned} \tag{3.3}$$

In the last equation we use $\operatorname{div}\sigma_h = f_h$, $\operatorname{div}\sigma_H = f_H$, $\operatorname{div}\xi_h = 0$, which are direct results of Hypothesis 1. Moreover,

$$\begin{aligned}
(f_h - f_H, v_h)_{L^2(\Omega)} &= (f_h - f_H, v_h - \mathcal{S}_H v_h)_{L^2(\Omega)} \\
&= \sum_{K \in \mathcal{T}_H \backslash \mathcal{T}_h} (f_h - f_H, v_h - \mathcal{S}_H v_h)_{L^2(K)}
\end{aligned} \tag{3.4}$$

By Hypothesis 3, a combination of (3.3) and (3.4) implies

$$\|\sigma_h - (\sigma_H - \xi_h)\|_{\mathcal{A}}^2 \lesssim \sum_{K \in \mathcal{T}_H \backslash \mathcal{T}_h} \|h_K(f - f_H)\|_{L^2(K)}^2. \tag{3.5}$$

Next, we analyze the second term on the right-hand side of (3.2). By the definition of ξ_h, and (2.3),

$$\|\xi_h\|_{\mathcal{A}}^2 = (\mathcal{A}\,\xi_h, \xi_h)_{L^2(\Omega)} = (\mathcal{A}\,\sigma_H, \xi_h)_{L^2(\Omega)} = (\mathcal{A}\,\sigma_H, \xi_h - \tau_H)_{L^2(\Omega)}, \tag{3.6}$$

for any $\tau_H \in Z_H(0)$. Since $\operatorname{div}\xi_h = 0$, it follows from Hypothesis 4 that there exists $\varphi_h \in H_h(\operatorname{curl}, \Omega)$ such that $\xi_h = \operatorname{curl}\varphi_h$. Hence the decomposition from (2.9) with $\tau_H = \operatorname{curl}\Pi_H\varphi_h \in \Sigma_H$ implies that there exist $\psi \in H(\operatorname{curl}, \Omega)$ and $\phi \in L^2(\Omega)$ such that

$$\xi_h - \tau_H = \operatorname{curl}(\varphi_h - \Pi_H\varphi_h) = \operatorname{curl}\psi + \phi, \tag{3.7}$$

with $(\mathcal{A}\,\sigma_H, \phi) = 0$ and ψ satisfying (2.10). A summary of (3.6), (3.7) and (2.10) leads to

$$\|\xi_h\|_{\mathcal{A}}^2 = (\mathcal{A}\,\sigma_H, \operatorname{curl}\psi)_{L^2(\Omega)} = (\mathcal{A}\,\sigma_H, \operatorname{curl}\psi)_{\tilde{\mathcal{R}}}. \tag{3.8}$$

We need to estimate the term on the right hand side of (3.8). For the 2D case, an integration by parts plus (2.10) yield

$$\begin{aligned}\|\xi_h\|_{\mathcal{A}}^2 =& -\sum_{K\in\tilde{\mathcal{R}}}\Big((\operatorname{rot}(\mathcal{A}\sigma_H),\psi)_{L^2(K)} + \sum_{E\in K\cap\mathcal{E}_H}\int_E \mathcal{A}\sigma_H\cdot t_E\psi ds\Big)\\ \lesssim& \left(\sum_{K\in\tilde{\mathcal{R}}}\Big(\|h_K\operatorname{rot}(\mathcal{A}\sigma_H)\|_{L^2(K)}^2 + \sum_{E\in K\cap\mathcal{E}_H}\|h_E^{1/2}[\mathcal{A}\sigma_H\cdot t_E]\|_{L^2(E)}^2\Big)\right)^{1/2}\|\xi_h\|_{L^2(\Omega)}.\end{aligned}$$

Since $\operatorname{div}\xi_h = 0$, it follows from (1.2) that

$$\|\xi_h\|_{L^2(\Omega)} \lesssim \|\xi_h\|_{\mathcal{A}},$$

which immediately implies

$$\|\xi_h\|_{\mathcal{A}}^2 \lesssim \sum_{K\in\tilde{\mathcal{R}}}\Big(\|h_K\operatorname{rot}(\mathcal{A}\sigma_H)\|_{L^2(K)}^2 + \sum_{E\in K\cap\mathcal{E}_H}\|h_E^{1/2}[\mathcal{A}\sigma_H\cdot t_E]\|_{L^2(E)}^2\Big). \tag{3.9}$$

For the 3D case, a similar argument shows

$$\begin{aligned}\|\xi_h\|_{\mathcal{A}}^2 =& -\sum_{K\in\tilde{\mathcal{R}}}\Big((\operatorname{curl}(\mathcal{A}\sigma_H),\psi)_{L^2(K)} + \sum_{E\in K\cap\mathcal{E}_H}\int_E \mathcal{A}\sigma_H\times\nu_E\psi ds\Big)\\ \lesssim& \left(\sum_{K\in\tilde{\mathcal{R}}}\Big(\|h_K\operatorname{curl}(\mathcal{A}\sigma_H)\|_{L^2(K)}^2 + \sum_{E\in K\cap\mathcal{E}_H}\|h_E^{1/2}[\mathcal{A}\sigma_H\times\nu_E]\|_{L^2(E)}^2\Big)\right)^{1/2}\|\xi_h\|_{\mathcal{A}}.\end{aligned}$$

Therefore,

$$\|\xi_h\|_{\mathcal{A}}^2 \lesssim \sum_{K\in\tilde{\mathcal{R}}}\Big(\|h_K\operatorname{curl}(\mathcal{A}\sigma_H)\|_{L^2(K)}^2 + \sum_{E\in K\cap\mathcal{E}_H}\|h_E^{1/2}[\mathcal{A}\sigma_H\times\nu_E]\|_{L^2(E)}^2\Big). \tag{3.10}$$

Finally, the desired result follows from (3.2), (3.5), (3.9) and (3.10). □

To establish quasi-orthogonality, we follow the idea of [22] to introduce the following problem: Find $(\tilde{\sigma}_h, \tilde{u}_h) \in \Sigma_h \times U_h$ such that

$$\begin{aligned}(\mathcal{A}\tilde{\sigma}_h, \tau_h)_{L^2(\Omega)} - (\operatorname{div}\tau_h, \tilde{u}_h)_{L^2(\Omega)} = 0, \text{ for any } \tau_h\in\Sigma_h,\\ (\operatorname{div}\tilde{\sigma}_h, v_h)_{L^2(\Omega)} - (f_H, v_h)_{L^2(\Omega)} = 0, \text{ for any } v_h\in U_h.\end{aligned} \tag{3.11}$$

Lemma 3.2. *Given $f \in L^2(\Omega)$, let (σ, u) be the solution of (1.1), (σ_h, u_h) and (σ_H, u_H) be the solutions to the discrete problem (2.1) over the nested triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$ respectively, and let $(\tilde{\sigma}_h, \tilde{u}_h)$ be the solution of (3.11). Then*

$$(\mathcal{A}(\sigma - \sigma_h), \tilde{\sigma}_h - \sigma_H)_{L^2(\Omega)} = 0, \tag{3.12}$$

$$\|\sigma_h - \tilde{\sigma}_h\|_{\mathcal{A}} \le \sqrt{C_0}\operatorname{osc}(f, \mathcal{T}_H\backslash\mathcal{T}_h). \tag{3.13}$$

Proof. By Hypothesis 1, the definitions of σ, σ_h, $\tilde{\sigma}_h$ and σ_H imply (3.12) directly via

$$\begin{aligned}(\mathcal{A}(\sigma-\sigma_h),\tilde{\sigma}_h-\sigma_H)_{L^2(\Omega)} &= (u-u_h,\operatorname{div}(\tilde{\sigma}_h-\sigma_H))_{L^2(\Omega)}\\ &= (u-u_h,f_H-f_H)_{L^2(\Omega)}=0,\end{aligned}$$

which shows (3.12).

Since by (2.3) $(\mathcal{A}\,\sigma_h,\tau_h)_{L^2(\Omega)}=0$ and $(\mathcal{A}\,\tilde{\sigma}_h,\tau_h)_{L^2(\Omega)}=0$ for any $\tau_h\in Z_h(0)$,

$$(\mathcal{A}(\sigma_h-\tilde{\sigma}_h),\tau_h)_{L^2(\Omega)}=0,\ \text{ for any } \tau_h\in Z_h(0), \tag{3.14}$$

which, along with the relations $\operatorname{div}\sigma_h=f_h$, $\operatorname{div}\tilde{\sigma}_h=f_H$ which follow from Hypothesis 1, the estimate (3.4) through Hypothesis 3, implies

$$\begin{aligned}\|\sigma_h-\tilde{\sigma}_h\|_{\mathcal{A}} &\lesssim \|\sigma_h-\tilde{\sigma}_h\|_{L^2(\Omega)/Z_h(0)}\\ &\lesssim \sup_{v_h\in U_h}\frac{(\operatorname{div}(\sigma_h-\tilde{\sigma}_h),v_h)_{L^2(\Omega)}}{\|v_h\|_{1,h}}\\ &= \sup_{v_h\in U_h}\frac{(f_h-f_H,v_h)_{L^2(\Omega)}}{\|v_h\|_{1,h}}\\ &\lesssim \Big(\sum_{K\in\mathcal{T}_H\backslash\mathcal{T}_h}\|h_K(f-f_H)\|^2_{L^2(K)}\Big)^{1/2}.\end{aligned} \tag{3.15}$$

This proves (3.13). □

Theorem 3.3. *Given $f\in L^2(\Omega)$, let (σ,u) be the solution of* (1.1)*, (σ_h,u_h) and (σ_H,u_H) be the solutions to the discrete problem* (2.1) *over the nested triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$ respectively. Then*

$$(\mathcal{A}(\sigma-\sigma_h),\sigma_h-\sigma_H)_{L^2(\Omega)}\le\sqrt{C_0}\|\sigma-\sigma_h\|_{\mathcal{A}}\operatorname{osc}(f,\mathcal{T}_H\backslash\mathcal{T}_h). \tag{3.16}$$

Thus, for any $\delta\ge 0$,

$$(1-\delta)\|\sigma-\sigma_h\|^2_{\mathcal{A}}\le\|\sigma-\sigma_H\|^2_{\mathcal{A}}-\|\sigma_h-\sigma_H\|^2_{\mathcal{A}}+\frac{C_0}{\delta}\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h). \tag{3.17}$$

Proof. We follow the idea of [22, Theorem 3.2]. Let $(\tilde{\sigma}_h,\tilde{u}_h)$ be the solution of (3.11). By Hypothesis 1 and Lemma 3.2,

$$(\mathcal{A}(\sigma-\sigma_h),\tilde{\sigma}_h-\sigma_H)_{L^2(\Omega)}=0.$$

Thus,

$$(\mathcal{A}(\sigma-\sigma_h),\sigma_h-\sigma_H)_{L^2(\Omega)}=(\mathcal{A}(\sigma-\sigma_h),\sigma_h-\tilde{\sigma}_h)_{L^2(\Omega)}\le\|\sigma-\sigma_h\|_{\mathcal{A}}\|\sigma_h-\tilde{\sigma}_h\|_{\mathcal{A}}. \tag{3.18}$$

Therefore, the estimate (3.16) follows from the inequality (3.13). By the identity $\sigma-\sigma_H=\sigma-\sigma_h+\sigma_h-\sigma_H$,

$$\|\sigma-\sigma_H\|^2_{\mathcal{A}}=\|\sigma-\sigma_h\|^2_{\mathcal{A}}+\|\sigma_h-\sigma_H\|^2_{\mathcal{A}}+2(\mathcal{A}(\sigma-\sigma_h),\sigma_h-\sigma_H).$$

In general, we use

$$\begin{aligned}\|\sigma-\sigma_H\|^2_{\mathcal{A}} &= \|\sigma-\sigma_h\|^2_{\mathcal{A}}+\|\sigma_h-\sigma_H\|^2_{\mathcal{A}}+2(\mathcal{A}(\sigma-\sigma_h),\sigma_h-\sigma_H)\\ &\ge \|\sigma-\sigma_h\|^2_{\mathcal{A}}+\|\sigma_h-\sigma_H\|^2_{\mathcal{A}}-2\sqrt{C_0}\|\sigma-\sigma_h\|_{\mathcal{A}}\operatorname{osc}(f,\mathcal{T}_H\backslash\mathcal{T}_h)\\ &\ge \|\sigma_h-\sigma_H\|^2_{\mathcal{A}}+(1-\delta)\|\sigma-\sigma_h\|^2_{\mathcal{A}}-\frac{C_0}{\delta}\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h)\end{aligned} \tag{3.19}$$

to prove (3.17). In the last step, we have used the Young inequality. □

4. Convergence and Optimality of AMFEM

In this section, we prove convergence and optimality of the adaptive mixed finite element methods. First we present the adaptive algorithms. In what follows, we replace the dependence on the actual mesh $\mathcal{T}$ by the iteration counter k.

Algorithm. Given an initial mesh $\mathcal{T}_0$ and a marking parameter $0 < \theta < 1$, set $k = 0$ and iterate

- Solve on $\mathcal{T}_k$, to get the solution σ_k.
- Compute the error estimator $\eta = \eta(\sigma_k, \mathcal{T}_k)$.
- Mark the minimal element set $\mathcal{M}_k$, such that

$$\eta^2(\sigma_k, \mathcal{M}_k) + \operatorname{osc}^2(f, \mathcal{M}_k) \geq \theta(\eta^2(\sigma_k, \mathcal{T}_k) + \operatorname{osc}^2(f, \mathcal{T}_k)).$$

- Refine each element $K \in \mathcal{M}_k$ by the newest vertex bisection, and make some necessary completeness to get a refined conforming mesh $\mathcal{T}_{k+1}$; $k = k + 1$.

4.1. Convergence.

We follow the arguments in [21, 22, 29] to prove convergence of the above adaptive algorithms.

Lemma 4.1. *Given $f \in L^2(\Omega)$, let (σ, u) be the solution of (1.1), (σ_k, u_k) and (σ_{k-1}, u_{k-1}) be the solutions to the discrete problem (2.1) over the nested triangulations $\mathcal{T}_k$ and $\mathcal{T}_{k-1}$ respectively. Then given any positive constant ϵ, there exist positive constants $0 < \rho < 1$ (depending on the dimension) and $\beta_2(\epsilon)$ such that*

$$\eta^2(\sigma_k, \mathcal{T}_k) \leq (1 + \epsilon)(\eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) - \rho\eta^2(\sigma_{k-1}, \mathcal{M}_{k-1})) + \frac{1}{\beta_2(\epsilon)}\|\sigma_k - \sigma_{k-1}\|^2_{\mathcal{A}}, \tag{4.1}$$

and

$$\operatorname{osc}^2(f, \mathcal{T}_k) \leq \operatorname{osc}^2(f, \mathcal{T}_{k-1}) - \rho \operatorname{osc}^2(f, \mathcal{T}_{k-1}\backslash\mathcal{T}_k). \tag{4.2}$$

Proof. By the definition of $\eta^2(\sigma_k, \mathcal{T}_k)$, $\eta^2(\sigma_{k-1}, \mathcal{T}_k)$, the trace theorem and the inverse inequality imply

$$|\eta(\sigma_k, \mathcal{T}_k) - \eta(\sigma_{k-1}, \mathcal{T}_k)| \lesssim \|\sigma_k - \sigma_{k-1}\|_{\mathcal{A}}.$$

An application of the Young inequality yields

$$\eta^2(\sigma_k, \mathcal{T}_k) \leq (1 + \epsilon)\eta^2(\sigma_{k-1}, \mathcal{T}_k) + \frac{1}{\beta_2(\epsilon)}\|\sigma_k - \sigma_{k-1}\|^2_{\mathcal{A}}. \tag{4.3}$$

Let $\mathcal{N}_k = \mathcal{T}_k\backslash\mathcal{T}_{k-1}$ be the set of the new elements in $\mathcal{T}_k$ but not in $\mathcal{T}_{k-1}$, and $\bar{\mathcal{M}}_{k-1} \subseteq \mathcal{T}_{k-1}$ be the set of the elements which are refined. Notice that $\mathcal{T}_{k-1}\backslash\bar{\mathcal{M}}_{k-1} = \mathcal{T}_k\backslash\mathcal{N}_k$. Given element $K \in \mathcal{N}_k$, consider its edge/face $E \in K \cap \mathcal{E}_k$. If E is in the interior of some element $T \in \mathcal{M}_{k-1}$, then $[\mathcal{A}\,\sigma_{k-1} \times \nu_E]|_E = 0$ since σ_{k-1} is a polynomial in K; otherwise, its measure is at most half of that of some edge/face of $T \in \mathcal{M}_{k-1}$ and thus

$$\eta^2(\sigma_{k-1}, \mathcal{N}_k) \leq 2^{-\frac{1}{d-1}}\eta^2(\sigma_{k-1}, \bar{\mathcal{M}}_{k-1}).$$

Since some more elements are refined for the conformity of the triangulation, $\mathcal{M}_{k-1} \subseteq \bar{\mathcal{M}}_{k-1}$. Therefore,

$$\begin{aligned} \eta^2(\sigma_{k-1}, \mathcal{T}_k) &= \eta^2(\sigma_{k-1}, \mathcal{N}_k) + \eta^2(\sigma_{k-1}, \mathcal{T}_k \backslash \mathcal{N}_k) \\ &\leq 2^{-\frac{1}{d-1}} \eta^2(\sigma_{k-1}, \bar{\mathcal{M}}_{k-1}) + \eta^2(\sigma_{k-1}, \mathcal{T}_{k-1} \backslash \bar{\mathcal{M}}_{k-1}) \\ &\leq \eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) - (1 - 2^{-\frac{1}{d-1}}) \eta^2(\sigma_{k-1}, \bar{\mathcal{M}}_{k-1}) \\ &\leq \eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) - (1 - 2^{-\frac{1}{d-1}}) \eta^2(\sigma_{k-1}, \mathcal{M}_{k-1}). \end{aligned}$$

This leads to

$$\eta^2(\sigma_{k-1}, \mathcal{T}_k) \leq \eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) - (1 - 2^{-\frac{1}{d-1}}) \eta^2(\sigma_{k-1}, \mathcal{M}_{k-1}). \tag{4.4}$$

With $\rho := 1 - 2^{-\frac{1}{d-1}}$, a combination of (4.3) and (4.4) proves (4.1). As for (4.2), it is an immediate result of the definition of the mesh size h_K. ☐

In the next theorem, we establish convergence of the adaptive methods.

Theorem 4.2. *Given $f \in L^2(\Omega)$, let (σ, u) be the solution of (1.1), (σ_k, u_k) and (σ_{k-1}, u_{k-1}) be the solutions to the discrete problem (2.1) over the nested triangulations $\mathcal{T}_k$ and $\mathcal{T}_{k-1}$ respectively. Then there exist positive constants $0 < \alpha < 1$, $\beta > 0$, $\gamma > 0$ such that*

$$\epsilon_k \leq \alpha \epsilon_{k-1},$$

where

$$\epsilon_k = \|\sigma - \sigma_k\|_{\mathcal{A}}^2 + \gamma \eta^2(\sigma_k, \mathcal{T}_k) + (\beta + \gamma) \operatorname{osc}^2(f, \mathcal{T}_k).$$

Proof. It follows from Theorem 3.3, (4.1) and (4.2) that

$$\begin{aligned} \epsilon_k \leq\ & \frac{1}{1-\delta} \|\sigma - \sigma_{k-1}\|_{\mathcal{A}}^2 - \frac{1}{1-\delta} \|\sigma_k - \sigma_{k-1}\|_{\mathcal{A}}^2 + \frac{C_0}{\delta(1-\delta)} \operatorname{osc}^2(f, \mathcal{T}_{k-1} \backslash \mathcal{T}_k) \\ & + \gamma(1+\epsilon)(\eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) - \rho \eta^2(\sigma_{k-1}, \mathcal{M}_{k-1})) + \frac{\gamma}{\beta_2(\epsilon)} \|\sigma_k - \sigma_{k-1}\|_{\mathcal{A}}^2 \\ & + (\beta + \gamma)(\operatorname{osc}^2(f, \mathcal{T}_{k-1}) - \rho \operatorname{osc}^2(f, \mathcal{T}_{k-1} \backslash \mathcal{T}_k)). \end{aligned}$$

Now the choice of $\gamma = \frac{\beta_2(\epsilon)}{1-\delta}$ and $\beta = \frac{C_0}{\rho\delta(1-\delta)}$ leads to

$$\begin{aligned} \epsilon_k \leq\ & \frac{1}{1-\delta} \|\sigma - \sigma_{k-1}\|_{\mathcal{A}}^2 + \beta \operatorname{osc}^2(f, \mathcal{T}_{k-1}) \\ & + \gamma(1+\epsilon)\big(\eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) + \operatorname{osc}^2(f, \mathcal{T}_{k-1})\big) \\ & - \gamma(1+\epsilon)\rho\big(\eta^2(\sigma_{k-1}, \mathcal{M}_{k-1}) + \operatorname{osc}^2(f, \mathcal{T}_{k-1} \backslash \mathcal{T}_k)\big). \end{aligned}$$

Since $\mathcal{M}_{k-1} \subset \mathcal{T}_{k-1} \backslash \mathcal{T}_k$, the marking strategy in adaptive Algorithm implies that

$$\eta^2(\sigma_{k-1}, \mathcal{M}_{k-1}) + \operatorname{osc}^2(f, \mathcal{T}_{k-1} \backslash \mathcal{T}_k) \geq \theta(\eta^2(\sigma_k, \mathcal{T}_{k-1}) + \operatorname{osc}^2(f, \mathcal{T}_{k-1})).$$

A substitution of this inequality into the previous one yields

$$\begin{aligned} \epsilon_k \leq\ & \frac{1}{1-\delta} \|\sigma - \sigma_{k-1}\|_{\mathcal{A}}^2 + \beta \operatorname{osc}^2(f, \mathcal{T}_{k-1}) \\ & + \gamma(1+\epsilon)(1 - \rho\theta)(\eta^2(\sigma_{k-1}, \mathcal{T}_{k-1}) + \operatorname{osc}^2(f, \mathcal{T}_{k-1})). \end{aligned}$$

By the definition of ϵ_{k-1}, we have for any $0<\alpha<1$,

$$\begin{aligned}\epsilon_k-\alpha\epsilon_{k-1} &\leq (\frac{1}{1-\delta}-\alpha)\|\sigma-\sigma_{k-1}\|_{\mathcal{A}}^2+\beta(1-\alpha)\operatorname{osc}^2(f,\mathcal{T}_{k-1})\\ &\quad+\gamma((1+\epsilon)(1-\rho\theta)-\alpha)(\eta^2(\sigma_{k-1},\mathcal{T}_{k-1})+\operatorname{osc}^2(f,\mathcal{T}_{k-1})).\end{aligned}$$

Hypothesis 5 states

$$\|\sigma-\sigma_{k-1}\|_{\mathcal{A}}^2\leq C_{Rel}(\eta^2(\sigma_{k-1},\mathcal{T}_{k-1})+\operatorname{osc}^2(f,\mathcal{T}_{k-1})).$$

This inequality plus the inequality give

$$\begin{aligned}\epsilon_k-\alpha\epsilon_{k-1} &\leq \left((\frac{1}{1-\delta}-\alpha)C_{Rel}+\gamma(1+\epsilon)(1-\rho\theta)-\gamma\alpha\right)\eta^2(\sigma_{k-1},\mathcal{T}_{k-1})\\ &\quad+\left((\frac{1}{1-\delta}-\alpha)C_{Rel}+\gamma(1+\epsilon)(1-\rho\theta)-\gamma\alpha+\beta(1-\alpha)\right)\operatorname{osc}^2(f,\mathcal{T}_{k-1}).\end{aligned}$$

To ensure $\epsilon_k-\alpha\epsilon_{k-1}\leq 0$, the factor α can be chosen such that

$$(\frac{1}{1-\delta}-\alpha)C_{Rel}+\gamma(1+\epsilon)(1-\rho\theta)-\gamma\alpha+\beta(1-\alpha)\leq 0,$$

which implies $\alpha=\frac{\beta+\gamma(1+\epsilon)(1-\rho\theta)+\frac{C_{Rel}}{1-\delta}}{\beta+\gamma+C_{Rel}}$ with $0<\delta<\frac{\gamma(\rho\theta-\epsilon(1-\rho\theta))}{\gamma(\rho\theta-\epsilon(1-\rho\theta))+C_{Rel}}$. □

4.2. **Optimality.** Let $\mathcal{T}_0$ be an initial quasi-uniform triangulation with $\#\mathcal{T}_0>2$, and let $\mathbb{T}_N$ be the set of all possible triangulations $\mathcal{T}$ which is generated from $\mathcal{T}_0$ with at most N elements more than $\mathcal{T}_0$. For $s>0$ we define the approximation class $\mathbb{A}_s$ as

$$\mathbb{A}_s:=\{(\sigma,f):|\sigma,f|_s<\infty,\ \text{with } |\sigma,f|_s:=\sup_{N>0}(N^s\inf_{\mathcal{T}\in\mathbb{T}_N}\inf_{\tau\in\Sigma_{\mathcal{T}}}\|\sigma-\tau\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}))\}.$$

Lemma 4.3. *Given a parameter*

$$\theta\in\left(0,\frac{\min\{C_{Eff},1\}}{C_{Drel}+\min\{C_{Eff},1\}+1}\right), \tag{4.5}$$

let (σ,u) *be the solution of* (1.1), (σ_h,u_h) *and* (σ_H,u_H) *be the solutions to the discrete problem* (2.1) *over* $\mathcal{T}_h$ *and* $\mathcal{T}_H$, *satisfying*

$$\|\sigma-\sigma_h\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_h)\leq\alpha'(\|\sigma-\sigma_H\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_H)) \tag{4.6}$$

with $0<\alpha'<\frac{\min\{C_{Eff},1\}-(\min\{C_{Eff},1\}+C_{Drel}+1)\theta}{\min\{C_{Eff},1\}+C_0}\in(0,1)$, *then it holds*

$$\theta(\eta^2(\sigma_H,\mathcal{T}_H)+\operatorname{osc}^2(f,\mathcal{T}_H))\leq\eta^2(\sigma_H,\tilde{\mathcal{R}})+\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h).$$

Proof. On one hand, from Theorem 3.1 it holds

$$\|\sigma_h-\sigma_H\|_{\mathcal{A}}^2\leq C_{Drel}(\eta^2(\sigma_H,\tilde{\mathcal{R}})+\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h)). \tag{4.7}$$

On the other hand, from Theorem 3.3 and the Young inequality it holds

$$\begin{aligned}2(\sigma-\sigma_h,\sigma_h-\sigma_H)_{\mathcal{A}} &\leq 2\sqrt{C_0}\|\sigma-\sigma_h\|_{\mathcal{A}}\operatorname{osc}(f,\mathcal{T}_H\backslash\mathcal{T}_h)\\ &\leq \frac{C_0}{\min\{C_{Eff},1\}}\|\sigma-\sigma_h\|_{\mathcal{A}}^2+\min\{C_{Eff},1\}\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h).\end{aligned}$$

This leads to

$$\begin{aligned}\|\sigma_h-\sigma_H\|_{\mathcal{A}}^2 &= \|\sigma-\sigma_H\|_{\mathcal{A}}^2-\|\sigma-\sigma_h\|_{\mathcal{A}}^2-2(\sigma-\sigma_h,\sigma_h-\sigma_H)_{\mathcal{A}}\\ &\geq (\|\sigma-\sigma_H\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_H))\\ &\quad -(1+\frac{C_0}{\min\{C_{Eff},1\}})(\|\sigma-\sigma_h\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_h))\\ &\quad -\operatorname{osc}^2(f,\mathcal{T}_H)+\operatorname{osc}^2(f,\mathcal{T}_h)-\min\{C_{Eff},1\}\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h).\end{aligned}$$

The condition (4.6), the lower bound in Hypothesis 5, and the relation

$$|\operatorname{osc}^2(f,\mathcal{T}_H)-\operatorname{osc}^2(f,\mathcal{T}_h)|\leq \operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h)$$

imply

$$\begin{aligned}\|\sigma_h-\sigma_H\|_{\mathcal{A}}^2 &\geq (\min\{C_{Eff},1\}-(\min\{C_{Eff},1\}+C_0)\alpha')\\ &\quad\times(\eta^2(\sigma_H,\mathcal{T}_H)+\operatorname{osc}^2(f,\mathcal{T}_H))\\ &\quad-(\min\{C_{Eff},1\}+1)\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h).\end{aligned}\tag{4.8}$$

A combination of (4.7) and (4.8) yields

$$\begin{aligned}&(C_{Drel}+\min\{C_{Eff},1\}+1)(\eta^2(\sigma_H,\tilde{\mathcal{R}})+\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h))\\ &\geq(\min\{C_{Eff},1\}-(\min\{C_{Eff},1\}+C_0)\alpha')(\eta^2(\sigma_H,\mathcal{T}_H)+\operatorname{osc}^2(f,\mathcal{T}_H)),\end{aligned}$$

from which, and the definition of α' and the restriction on θ, we obtain the desired result. □

Lemma 4.4. *Let (σ,u) be the solution of (1.1), (σ_h,u_h) and (σ_H,u_H) be the solutions to the discrete problem (2.1) over $\mathcal{T}_h$ and $\mathcal{T}_H$, there exists a constant $C_1>0$ such that*

$$\|\sigma-\sigma_h\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_h)\leq C_1(\|\sigma-\sigma_H\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_H)).\tag{4.9}$$

Proof. From (3.17) and (4.2), for any $0<\delta<1$, it holds

$$\begin{aligned}&\|\sigma-\sigma_h\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_h)\\ &\leq\frac{1}{1-\delta}\|\sigma-\sigma_H\|_{\mathcal{A}}^2+\frac{C_0}{\delta(1-\delta)}\operatorname{osc}^2(f,\mathcal{T}_H\backslash\mathcal{T}_h)+\operatorname{osc}^2(f,\mathcal{T}_h)\\ &\leq\frac{1}{1-\delta}\|\sigma-\sigma_H\|_{\mathcal{A}}^2+\left(\frac{C_0}{\delta(1-\delta)}+1\right)\operatorname{osc}^2(f,\mathcal{T}_H),\end{aligned}$$

which implies the desired result. □

Theorem 4.5. *Let $\mathcal{M}_k$ be a set of marked elements with minimal cardinality, (σ,u) the solution of (1.1), and $(\mathcal{T}_k,\Sigma_k,\sigma_k,u_k)$ the sequence of triangulations, finite element spaces and discrete solutions produced by the adaptive finite element methods with the marking parameter θ in Lemma 4.3. It holds that*

$$\#\mathcal{M}_k\leq(\alpha')^{-\frac{1}{s}}|u,f|_s^{\frac{1}{s}}C_1^{\frac{1}{s}}C_2(\|\sigma-\sigma_k\|_{\mathcal{A}}^2+\operatorname{osc}^2(f,\mathcal{T}_k))^{-\frac{1}{s}},$$

where α' is defined in Lemma 4.3, C_1 in Lemma 4.4, and C_2 only depends on the shape regularity of $\mathcal{T}_0$.

Proof. We set $\varepsilon = \alpha' C_1^{-1}(\|\sigma-\sigma_k\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_k))$ where α' is from Lemma 4.3, and C_1 from Lemma 4.4. Since $(\sigma, f)\in\mathbb{A}_s$, there exist a refinement of $\mathcal{T}_0$, say, $\mathcal{T}_\varepsilon$, and $\sigma_\varepsilon\in\Sigma_{\mathcal{T}_\varepsilon}$ such that

$$\#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \le |\sigma, f|_s^{\frac{1}{s}}\varepsilon^{-\frac{1}{s}},$$
$$\|\sigma-\sigma_\varepsilon\|_{\mathcal{A}}^2 + \mathrm{osc}^2(f,\mathcal{T}_\varepsilon)\le\varepsilon.$$

Let $\mathcal{T}_*$ be the overlay of $\mathcal{T}_\varepsilon$ and $\mathcal{T}_k$, and (σ_*, u_*) be the corresponding discrete solution on $\mathcal{T}_*$. Since $\mathcal{T}_*$ is a refinement of $\mathcal{T}_\varepsilon$, it follows from Lemma 4.4 that

$$\|\sigma-\sigma_*\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_*)\le C_1(\|\sigma-\sigma_\varepsilon\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_\varepsilon))\le C_1\varepsilon=\alpha'(\|\sigma-\sigma_k\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_k)).$$

From Lemma 4.3 it holds

$$\theta(\eta^2(\sigma_k,\mathcal{T}_k)+\mathrm{osc}^2(f,\mathcal{T}_k))\le\eta^2(\sigma_k,\widetilde{\mathcal{T}_k\backslash\mathcal{T}_*})+\mathrm{osc}^2(f,\mathcal{T}_k\backslash\mathcal{T}_*),$$

here $\widetilde{\mathcal{T}_k\backslash\mathcal{T}_*}$ is similarly defined as $\tilde{\mathcal{R}}$. Note that the marking step in the adaptive Algorithm with θ chooses a subset of $\mathcal{M}_k\subset\mathcal{T}_k$ with minimal cardinality so that the same property holds. Therefore, there exists a constant C_2 depending on the shape regularity of $\mathcal{T}_0$ such that

$$\#\mathcal{M}_k\le\#\widetilde{\mathcal{T}_k\backslash\mathcal{T}_*}\le C_2(\#\mathcal{T}_*-\#\mathcal{T}_k)\le C_2(\#\mathcal{T}_\varepsilon-\#\mathcal{T}_0).$$

By the definition of ε, a combination of the above inequalities shows

$$\#\mathcal{M}_k\le(\alpha')^{-\frac{1}{s}}|\sigma,f|_s^{\frac{1}{s}}C_1^{\frac{1}{s}}C_2(\|\sigma-\sigma_k\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_k))^{-\frac{1}{s}}.$$

□

Theorem 4.6. *Let $\mathcal{M}_k$ be a set of marked elements with minimal cardinality, (σ,u) the solution of* (1.1)*, and $(\mathcal{T}_k,\Sigma_k,\sigma_k,u_k)$ the sequence of triangulations, finite element spaces and discrete solutions produced by the adaptive finite element methods with the marking parameter θ in Lemma 4.3. Then it holds that*

$$\|\sigma-\sigma_N\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_N)\lesssim|\sigma,f|_s(\#\mathcal{T}_N-\#\mathcal{T}_0)^{-s},\quad \textit{for } (\sigma,f)\in\mathbb{A}_s.$$

Proof. Let $\mu=(\alpha')^{-\frac{1}{s}}|u,f|_s^{\frac{1}{s}}C_1^{\frac{1}{s}}C_2$. We use the result $\#\mathcal{T}_k-\#\mathcal{T}_0\lesssim\sum_{j=0}^{k-1}\#\mathcal{M}_j$ from [39], and the upper bound of $\#\mathcal{M}_j$ in Theorem 4.5, to obtain that

$$\#\mathcal{T}_N-\#\mathcal{T}_0\lesssim\sum_{j=0}^{N-1}\#\mathcal{M}_j\le\sum_{j=0}^{N-1}\mu(\|\sigma-\sigma_j\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_j))^{-\frac{1}{s}}.$$

From the convergence result in Theorem 4.2 we have $\epsilon_N\le\alpha^{N-j}\epsilon_j$ for any $0\le j\le N-1$, which, along with the fact $\epsilon_j\approx\|\sigma-\sigma_j\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_j)$, implies

$$\#\mathcal{T}_N-\#\mathcal{T}_0\lesssim\mu(\|\sigma-\sigma_N\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_N))^{-\frac{1}{s}}\sum_{j=0}^{N-1}\alpha^{\frac{j}{s}}.$$

Since $\alpha<1$, the term $\sum_{j=0}^{N-1}\alpha^{\frac{j}{s}}$ is bounded. The definition of μ leads to

$$\|\sigma-\sigma_N\|_{\mathcal{A}}^2+\mathrm{osc}^2(f,\mathcal{T}_N)\lesssim|u,f|_s(\#\mathcal{T}_N-\#\mathcal{T}_0)^{-s},\quad\text{for } (u,f)\in\mathbb{A}_s.$$

□

5. Applications

In this section, we present two examples which satisfy these five hypotheses. The first example is the mixed finite element of the Poisson equation; the second one is the mixed finite element of the Stokes problem within the pseudostress-velocity formulation. For both the 2D and 3D, we prove that the Raviart–Thomas and Brezzi–Douglas–Marini elements satisfy Hypotheses 1–5 in Section 2. Hence the corresponding adaptive algorithms converge at the optimal rate in the nonlinear approximation sense.

In the sequel, we use superscript P to denote the subspace or operator for the Poisson problem, and S to denote the subspace or operator for the Stokes problem.

5.1. **The Poisson problem.**

The Raviart–Thomas element spaces [10] are defined for $k \geq 0$ by

$$RT_h^P = \Sigma_{h,k}^P \times U_{h,k}^P,$$

where

$$\Sigma_{h,k}^P := \{\tau \in H(\mathrm{div}, \Omega; \mathbb{R}^d) : \tau|_K \in P_k(K)^d + xP_k(K), \forall K \in \mathcal{T}_h\},$$

and

$$U_{h,k}^P := \{v \in L^2(\Omega; \mathbb{R}) : v|_K \in P_k(K), \forall K \in \mathcal{T}_h\}.$$

Here $P_k(K)$ denotes the space of polynomials of degree at most k over K.

The Brezzi–Douglas–Marini element spaces [10] are defined for $k \geq 0$ by

$$BDM_h^P = \Sigma_{h,k}^P \times U_{h,k}^P,$$

where

$$\Sigma_{h,k}^P := \{\tau \in H(\mathrm{div}, \Omega; \mathbb{R}^d) : \tau|_K \in P_{k+1}(K)^d, \forall K \in \mathcal{T}_h\},$$

and

$$U_{h,k}^P := \{v \in L^2(\Omega; \mathbb{R}) : v|_K \in P_k(K), \forall K \in \mathcal{T}_h\}.$$

Theorem 5.1. *For the Raviart–Thomas and Brezzi–Douglas–Marini elements of the Poisson problem, Hypotheses 1–3 and 5 hold.*

Proof. Hypothesis 1 is an immediate result of the definitions of finite element subspaces, while Hypotheses 2, 3, and 5 were proved in [32, Lemma 2.1], [29, Lemma 2.8], [29, Theorem 2.1, 2.2], respectively. □

In order to check Hypothesis 4, we need the following two spaces:

$$S_{h,k} := \{\psi \in H^1(\Omega; \mathbb{R}) : \psi|_K \in P_k(K), \forall K \in \mathcal{T}_h\}, \quad d = 2,$$

$$ND_{h,k} := \{\psi \in H(\mathrm{curl}, \Omega; \mathbb{R}^3) : \psi|_K \in ND_k(K), \forall K \in \mathcal{T}_h\}, \quad d = 3,$$

where $ND_k(K) := P_{k-1}(K)^3 \oplus \{v \in \tilde{P}_k(K)^3, v \cdot x = 0, \forall x \in K\}$, with $\tilde{P}_k(K)$ being the spaces of homogeneous polynomials of degree k over K.

Lemma 5.2. [43, Theorem 4.1] [29, Lemma 2.10] *There exists a quasi-interpolation operator* $\mathcal{P}_{H,k}$: $ND_{h,k} \to ND_{H,k}$ *such that for any* $\varphi \in ND_{h,k}$*,*

$$\mathcal{P}_{H,k}\varphi|_{\mathcal{T}_H\backslash\tilde{\mathcal{R}}} = \varphi|_{\mathcal{T}_H\backslash\tilde{\mathcal{R}}}, \tag{5.1}$$

$$\|\operatorname{curl}\mathcal{P}_{H,k}\varphi\|_{L^2(\Omega)} \lesssim \|\operatorname{curl}\varphi\|_{L^2(\Omega)}. \tag{5.2}$$

Lemma 5.3. [37, Theorem 1] *There exists an operator* $\mathcal{I}_H : H(\operatorname{curl},\Omega;\mathbb{R}^3) \to ND_{H,1}$ *with the following properties: For every* $\varphi \in H(\operatorname{curl},\Omega;\mathbb{R}^3)$*, there exist* $\psi \in H^1(\Omega;\mathbb{R}^3)$ *and* $w \in H^1(\Omega;\mathbb{R})$ *such that*

$$\varphi - \mathcal{I}_H\varphi = \psi + \operatorname{grad} w,$$

and

$$h_K^{-1}\|w\|_{L^2(K)} + \|\operatorname{grad} w\|_{L^2(K)} \lesssim \|\varphi\|_{L^2(\Omega_K)},$$
$$h_K^{-1}\|\psi\|_{L^2(K)} + \|\operatorname{grad}\psi\|_{L^2(K)} \lesssim \|\operatorname{curl}\varphi\|_{L^2(\Omega_K)}.$$

Now we are ready to present the following theorem.

Theorem 5.4. *For the Raviart–Thomas and Brezzi–Douglas–Marini elements of the Poisson problem, Hypothesis 4 holds.*

Proof. For the 2D case, $H(\operatorname{curl},\Omega) = H^1(\Omega;\mathbb{R})$. Let $H_h^P(\operatorname{curl},\Omega) = S_{h,k}$, and Π_H^P be the Scott-Zhang interpolation operator [38]. Then (2.8) is true. Let $\psi = \varphi_h - \Pi_H^P\varphi_h$, and $\phi = 0$, which implies that the decomposition (2.9), the estimates (2.10) hold.

For the 3D case, let $H_h^P(\operatorname{curl},\Omega) = ND_{h,k}$, and $\Pi_H^P = \mathcal{I}_H + \mathcal{P}_{H,k} - \mathcal{I}_H\mathcal{P}_{H,k}$ where the operate $\mathcal{I}_H$ is from Lemma 5.3, and the operator $\mathcal{P}_{H,k}$ is from Lemma 5.2. Then (2.8) is true. In addition, from Lemma 5.3, there exist $\psi \in H^1(\Omega;\mathbb{R}^3)$ and $w \in H^1(\Omega;\mathbb{R})$ such that

$$\varphi_h - \Pi_H^P\varphi_h = (I - \mathcal{I}_H)(I - \mathcal{P}_{H,k})\varphi_h = \psi + \operatorname{grad} w, \tag{5.3}$$

and

$$h_K^{-2}\|\psi\|_{L^2(K)}^2 + \|\operatorname{grad}\psi\|_{L^2(K)}^2 \lesssim \|\operatorname{curl}(I - \mathcal{P}_{H,k})\varphi_h\|_{L^2(\Omega_K)}^2. \tag{5.4}$$

From (5.3), (5.4), the trace theorem, and Lemma 5.2, we obtain (2.9) and (2.10) with $\phi = 0$. □

5.2. **The Stokes problem.** The finite element spaces for the Stokes problem are defined rowwise based on those for the Poisson problem case, with an additional restriction that the mean of the trace of the stress vanishes, namely,

$$\Sigma_{h,k}^S := \left\{\tau \in (\Sigma_{h,k}^P)^d | \int_\Omega \operatorname{tr}\tau dx = 0\right\}, \text{ and } U_{h,k}^S := (U_{h,k}^P)^d.$$

Theorem 5.5. *For the Raviart–Thomas and Brezzi–Douglas–Marini elements of the Stokes problem, Hypotheses 1–3 and 5 hold.*

Proof. The proofs of Hypothesis 1 and 3 are similar to those for the Poisson problem. The proof of Hypothesis 5 can be found in [19, Theorem 5.4] for the 2D case while the proof of the 3D case is similar. In order to show Hypothesis 2, we need to handle the additional restriction on the mean of the trace of the stress. In fact, given $v_h = (v_1, \cdots, v_d)^T \in U^S_{h,k}$ with $v_i \in U^P_{h,k}$ $(i = 1, \cdots, d)$, there exists $\tau_i \in \Sigma^P_{h,k}$ such that

$$\|v_i\|_{1,h} \lesssim \frac{(\operatorname{div} \tau_i, v_i)_{L^2(\Omega)}}{\|\tau_i\|_{L^2(\Omega)}}.$$

Let $\tau_h = (\tau_1, \cdots, \tau_d)^T$, and define

$$\tilde{\tau}_h := \tau_h - \frac{\int_\Omega \operatorname{tr} \tau_h dx}{d|\Omega|} I_{d\times d},$$

then $\tilde{\tau}_h \in \Sigma^S_{h,k}$ and

$$(\operatorname{div} \tilde{\tau}_h, v_h)_{L^2(\Omega)} = (\operatorname{div} \tau_h, v_h)_{L^2(\Omega)}.$$

Since

$$\|\tilde{\tau}_h\|^2_{L^2(\Omega)} = \|\tau_h\|^2_{L^2(\Omega)} - \frac{(\int_\Omega \operatorname{tr} \tau_h dx)^2}{d|\Omega|} \le \|\tau_h\|^2_{L^2(\Omega)},$$

we immediately have

$$\|v_h\|_{1,h} \lesssim \frac{(\operatorname{div} \tau_h, v_h)_{L^2(\Omega)}}{\|\tau_h\|_{L^2(\Omega)}} \le \frac{(\operatorname{div} \tilde{\tau}_h, v_h)_{L^2(\Omega)}}{\|\tilde{\tau}_h\|_{L^2(\Omega)}}.$$

This proves Hypothesis 2. □

Theorem 5.6. *For the Raviart–Thomas and Brezzi–Douglas–Marini elements of the Stokes problem, Hypothesis 4 holds.*

Proof. For this case, the finite element space $H^S_h(\operatorname{curl}, \Omega) := (H^P_h(\operatorname{curl}, \Omega))^d$. For $\varphi_h = (\varphi_1, \cdots, \varphi_d)^T \in H^S_h(\operatorname{curl}, \Omega)$, define $\tilde{\Pi}^S_H \varphi_h = (\Pi^P_H \varphi_1, \cdots, \Pi^P_H \varphi_d)^T$, and

$$\Pi^S_H \varphi_h := \tilde{\Pi}^S_H \varphi_h - \frac{\int_\Omega \operatorname{tr}(\operatorname{curl} \tilde{\Pi}^S_H \varphi_h) dx}{d|\Omega|} \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}, \quad d = 2,$$

$$\Pi^S_H \varphi_h := \tilde{\Pi}^S_H \varphi_h - \frac{\int_\Omega \operatorname{tr}(\operatorname{curl} \tilde{\Pi}^S_H \varphi_h) dx}{2d|\Omega|} \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}, \quad d = 3.$$

This yieds

$$\operatorname{curl} \Pi^S_H \varphi_h = \operatorname{curl} \tilde{\Pi}^S_H \varphi_h - \frac{\int_\Omega \operatorname{tr}(\operatorname{curl} \tilde{\Pi}^S_H \varphi_h) dx}{d|\Omega|} I_{d\times d} \text{ and } \int_\Omega \operatorname{tr}(\operatorname{curl} \Pi^S_H \varphi_h) dx = 0,$$

and so $\operatorname{curl} \Pi^S_H \varphi_h \in \Sigma^S_H$.

For any $\varphi_i \in H^P_h(\operatorname{curl}, \Omega)$, from Theorem 5.4 there exists $\psi_i \in H^1(\Omega)$ such that

$$\operatorname{curl}(\varphi_i - \Pi^P_H \varphi_i) = \operatorname{curl} \psi_i,$$

and ψ_i satisfy the condition of (2.10). Let $\psi := (\psi_1, \cdots, \psi_d)^T$. It holds that

$$\operatorname{curl}(\varphi_h - \Pi^S_H \varphi_h) = \operatorname{curl} \psi + \frac{\int_\Omega \operatorname{tr}(\operatorname{curl} \tilde{\Pi}^S_H \varphi_h) dx}{d|\Omega|} I_{d\times d}.$$

Hence the condition (2.10) is satisfied with $\phi := \frac{\int_\Omega \operatorname{tr}(\operatorname{curl} \tilde{\Pi}^S_H \varphi_h) dx}{d|\Omega|} I_{d\times d}$. In addition, $(\mathcal{A}\, \sigma_H, \phi)_{L^2(\Omega)} = 0$.

□

References

[1] M. Ainsworth. A posteriori error estimation for lowest order Raviart-Thomas mixed finite elements. *SIAM Journal on Scientific Computing*, 30(1):189–204, 2007.

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

[3] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM Journal on Numerical Analysis*, 19(4):742–760, 1982.

[4] D. N. Arnold and R. S. Falk. A new mixed formulation for elasticity. *Numerische Mathematik*, 53(1-2):13–30, 1988.

[5] I. Babuška and M. Vogelius. Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numerische Mathematik*, 44(1):75–102, 1984.

[6] R. Becker and S. Mao. An optimally convergent adaptive mixed finite element method. *Numerische Mathematik*, 111(1):35–54, 2008.

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

[8] D. Braess and R. Verfürth. A posteriori error estimators for the Raviart-Thomas element. *SIAM Journal on Numerical Analysis*, 33(6):2431–2444, 1996.

[9] J. H. Brandts. Superconvergence and a posteriori error estimation for triangular mixed finite elements. *Numerische Mathematik*, 68(3):311–324, 1994.

[10] F. Brezzi and M. Fortin. Mixed and hybrid finite element methods. 1991.

[11] Z. Cai, B. Lee, and P. Wang. Least-squares methods for incompressible newtonian fluid flow: Linear stationary problems. *SIAM Journal on Numerical Analysis*, 42(2):843–859, 2004.

[12] Z. Cai, C. Tong, P. S. Vassilevski, and C. Wang. Mixed finite element methods for incompressible flow: Stationary Stokes equations. *Numerical Methods for Partial Differential Equations*, 26(4):957–978, 2010.

[13] C. Carstensen. A posteriori error estimate for the mixed finite element method. *Mathematics of Computation of the American Mathematical Society*, 66(218):465–476, 1997.

[14] C. Carstensen. A unifying theory of a posteriori finite element error control. *Numerische Mathematik*, 100(4):617–637, 2005.

[15] C. Carstensen, M. Eigel, R. H. Hoppe, and C. Löbhard. A review of unified a posteriori finite element error control. *Numerical Mathematics: Theory, Methods and Applications*, 5(4):509–558, 2012.

[16] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Computers & Mathematics with Applications*, 67(6):1195–1253, 2014.

[17] C. Carstensen, D. Gallistl, and M. Schedensack. Quasi-optimal adaptive pseudostress approximation of the Stokes equations. *SIAM Journal on Numerical Analysis*, 51(3):1715–1734, 2013.

[18] C. Carstensen and R. Hoppe. Error reduction and convergence for an adaptive mixed finite element method. *Mathematics of computation*, 75(255):1033–1042, 2006.

[19] C. Carstensen, D. Kim, and E. J. Park. A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem. *SIAM Journal on Numerical Analysis*, 49(6):2501–2523, 2011.

[20] C. Carstensen and H. Rabus. An optimal adaptive mixed finite element method. *Mathematics of Computation*, 80(274):649–667, 2011.

[21] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM Journal on Numerical Analysis*, 46(5):2524–2550, 2008.

[22] L. Chen, M. Holst, and J. Xu. Convergence and optimality of adaptive mixed finite element methods. *Mathematics of Computation*, 78(265):35–53, 2009.

[23] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33(3):1106–1124, 1996.

[24] S. Du and X. Xie. Error reduction, convergence and optimality for adaptive mixed finite element methods for diffusion equations. *Journal of Computational Mathematics*, 30(5), 2012.

[25] S. Du and X. Xie. On residual-based a posteriori error estimators for lowest-order Raviart-Thomas element approximation to convection-diffusion-reaction equations. *Journal of Computational Mathematics*, 32(5):522–546, 2014.

[26] R. H. 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, 1997.

[27] J. Hu, Z. Shi, and J. Xu. Convergence and optimality of the adaptive Morley element method. *Numerische Mathematik*, 121(4):731–752, 2012.

[28] J. Hu and J. Xu. Convergence of adaptive conforming and nonconforming finite element methods for the perturbed Stokes equation. Technical report, Research Report. School of Mathematical Sciences and Institute of Mathematics, Peking University, 2007.

[29] J. Huang and Y. Xu. Convergence and complexity of arbitrary order adaptive mixed element methods for the Poisson equation. *Science China Mathematics*, 55(5):1083–1098, 2012.

[30] K.-Y. Kim. Guaranteed a posteriori error estimator for mixed finite element methods of elliptic problems. *Applied Mathematics and Computation*, 218(24):11820–11831, 2012.

[31] M. G. Larson and A. Målqvist. A posteriori error estimates for mixed finite element approximations of elliptic problems. *Numerische Mathematik*, 108(3):487–500, 2008.

[32] C. Lovadina and R. Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Mathematics of Computation*, 75(256):1659–1674, 2006.

[33] K. Mekchay and R. H. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM Journal on Numerical Analysis*, 43(5):1803–1827, 2005.

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

[35] P. Morin, R. Nochetto, and K. Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Mathematics of Computation*, 72(243):1067–1097, 2003.

[36] P. Morin, K. G. Siebert, and A. Veeser. A basic convergence result for conforming adaptive finite elements. *Mathematical Models and Methods in Applied Sciences*, 18(05):707–737, 2008.

[37] J. Schöberl. A posteriori error estimates for Maxwell equations. *Mathematics of Computation*, 77(262):633–649, 2008.

[38] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54(190):483–493, 1990.

[39] R. Stevenson. Optimality of a standard adaptive finite element method. *Foundations of Computational Mathematics*, 7(2):245–269, 2007.

[40] M. Vohralík. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM Journal on Numerical Analysis*, 45(4):1570–1599, 2007.

[41] M. F. Wheeler and I. Yotov. A posteriori error estimates for the mortar mixed finite element method. *SIAM Journal on Numerical Analysis*, 43(3):1021–1042, 2005.

[42] B. Wohlmuth and R. Hoppe. A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements. *Mathematics of Computation of the American Mathematical Society*, 68(228):1347–1378, 1999.

[43] L. Zhong, L. Chen, S. Shu, G. Wittum, and J. Xu. Convergence and optimality of adaptive edge finite element methods for time-harmonic Maxwell equations. *Mathematics of Computation*, 81(278):623–642, 2012.