

## QUASI-OPTIMAL ADAPTIVE HYBRIDIZED MIXED FINITE ELEMENT METHODS FOR LINEAR ELASTICITY

YUWEN LI\* 

**Abstract.** For the planar Navier–Lamé equation in mixed form with symmetric stress tensors, we prove the uniform quasi-optimal convergence of an adaptive method based on the hybridized mixed finite element proposed in Gong *et al.* [*Numer. Math.* **141** (2019) 569–604]. The main ingredients in the analysis consist of a discrete *a posteriori* upper bound and a quasi-orthogonality result for the stress field under the mixed boundary condition. Compared with existing adaptive methods, the proposed adaptive algorithm could be directly applied to the traction boundary condition and be easily implemented.

**Mathematics Subject Classification.** 65N12, 65N15, 65N30, 65N50.

Received March 28, 2021. Accepted August 18, 2021.

### 1. INTRODUCTION

Adaptive finite element methods for numerical solutions of partial differential equations has been an active research area since 1980s. Using a sequence of self-adapted graded meshes, adaptive methods can achieve quasi-optimal convergence rate even for problems with singularity arising from, *e.g.*, irregular data or domains with nonsmooth boundary. Convergence and optimality analysis of adaptive methods for symmetric and positive-definite elliptic problems has now reached maturity, see, *e.g.*, [7, 16, 21, 22, 38, 39, 41] and references therein.

An important model problem in linear elasticity is the Navier–Lamé equation, which could be discretized by primal methods and mixed methods. Adaptive mesh refinement based on *a posteriori* error indicators is essential to deal with nonsmooth boundaries of elastic bodies in practice. For conforming elasticity elements, a robust error estimator could be found in [10]. In [13], a quasi-optimal nonconforming adaptive Crouzeix–Raviart element method in primal form was developed under the pure displacement boundary condition.

Compared with primal methods, mixed methods could easily handle the traction boundary condition and is more natural from a viewpoint of solid mechanics. Conservation of angular momentum is implied by the symmetry of stress tensors of elasticity equations in mixed form. However, mixed methods with strongly imposed symmetry usually leads to higher order polynomial shape functions [2, 4, 28, 30] and *a priori* error estimates relying on high solution regularity. In such situations, adaptivity is of great importance, see, *e.g.*, [11, 15, 20, 34, 35] for *a posteriori* error estimates of adaptive mixed finite element methods (AMFEMs) in linear elasticity.

For second order elliptic equations in mixed form, theoretical analysis of AMFEMs is extensive, see, *e.g.*, [6, 12, 18, 23, 27, 31–33]. The optimality result of adaptive mixed methods for elasticity equations seems limited

---

*Keywords and phrases.* linear elasticity, mixed finite element, hybridization, convergence, quasi-optimality.

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA.

\*Corresponding author; [yuwenli925@gmail.com](mailto:yuwenli925@gmail.com)

in the literature. One reason is that most finite elements for discretizing the symmetric stress tensors require  $C^0$  vertex continuity. As a result, stress finite element spaces on nested meshes are not nested as spaces. Recently the work [29] develops a quasi-optimal AMFEM for mixed elasticity, based on a modified Hu–Zhang mixed element [30] enriched carefully at each vertex on nested meshes. In the meantime, Gong *et al.* [25] presents a hybridized mixed method for elasticity using complete piecewise polynomial stress space without any vertex degrees of freedom. In this paper, we shall adopt the hybridization strategy in [25] and develop a quasi-optimal adaptive mixed method for planar linear elasticity, see (1.4) and Algorithm 4.1. Without specific treatment, our AMFEM could directly be applied to the (pure) traction boundary condition. Another advantage of this AMFEM is its easy implementation because no explicit continuous local basis is needed in hybridization, see Section 6.

The proposed AMFEM is designed to reduce the stress error  $\|\sigma - \sigma_h\|_A$  with a convergence rate free of volumetric locking. The framework of our analysis is similar to the convergence analysis of AMFEMs for Poisson's equation [18, 31]. However, the  $C^1$  nodal space  $W_h$  used in our analysis (see Lem. 2.1) is much more complicated than  $C^0$  nodal spaces and the well-known  $C^1$  Argyris spaces. As a consequence, the regularized local interpolation onto  $W_h$  is rather involved, especially when tailored to respect mixed boundary conditions. In addition, we present a detailed construction of the discrete *a posteriori* upper bound in Theorem 3.3 for the stress error under general boundary conditions, which seems missing in the literature.

In the rest of this section, we introduce the continuous and discrete mixed formulations of the linear elasticity equation in  $\mathbb{R}^2$ . Let  $\Omega$  be a simply connected polygonal domain. Let  $\sigma$  and  $u$  denote the stress and displacement fields produced by a body force acting on a linearly elastic body that occupies the region  $\Omega \subset \mathbb{R}^2$ . Then  $u$  takes value in  $\mathbb{R}^2$  and  $\sigma$  takes value in  $\mathbb{S}$ , the space of symmetric  $2 \times 2$  matrices. Given Lamé constants  $\mu > 0$ ,  $\lambda > 0$ , define

$$\begin{aligned}\varepsilon(u) &= \frac{1}{2}(\nabla u + (\nabla u)^\top), \\ \mathbb{C}\sigma &= 2\mu\sigma + \lambda \operatorname{tr}(\sigma)\delta,\end{aligned}$$

where  $\operatorname{tr}$  denotes the trace of square matrices, and  $\delta$  is the  $2 \times 2$  identity matrix. The Navier–Lamé equation for planar elasticity reads

$$\operatorname{div}(\mathbb{C}\varepsilon(u)) = f, \tag{1.1}$$

where  $\operatorname{div}$  is the divergence operator applied to each row of  $\mathbb{C}\varepsilon(u)$ . Given  $\tau \in L^2(\Omega, \mathbb{S})$ , the compliance tensor is defined as

$$\mathbb{A}\tau = \mathbb{C}^{-1}\tau = \frac{1}{2\mu} \left( \tau - \frac{\lambda}{2\mu + 2\lambda} (\operatorname{tr} \tau)\delta \right).$$

Let  $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$  with relatively open subsets  $\Gamma_D$ ,  $\Gamma_N$ , and  $\Gamma_D \cap \Gamma_N = \emptyset$ . The part  $\Gamma_N = \cup_{j=1}^J \Gamma_j$  is the disjoint union of several connected components  $\{\Gamma_j\}_{j=1}^J$ . Let  $n$  be the outward unit normal to  $\partial\Omega$ . We consider the mixed formulation of (1.1) under the mixed boundary condition

$$\begin{aligned}\mathbb{A}\sigma &= \varepsilon(u), \\ \operatorname{div} \sigma &= f, \\ u &= g_D \text{ on } \Gamma_D, \\ \sigma n &= g_N \text{ on } \Gamma_N.\end{aligned} \tag{1.2}$$

Let  $\mathcal{RM}$  be the space of rigid body motions

$$\mathcal{RM} = \left\{ (c_1, c_2)^\top + c_3(-x_2, x_1)^\top : c_1, c_2, c_3 \in \mathbb{R} \right\}.$$

If  $\Gamma_D = \emptyset$ , the load  $f$  in (1.2) is required to satisfy the compatibility condition

$$\int_{\Omega} f \cdot v \, dx = 0, \quad \forall v \in \mathcal{RM}.$$

Given a vector space  $\mathbb{V}$ , let  $L^2(\Omega, \mathbb{V})$  denote the space of  $\mathbb{V}$ -valued  $L^2$ -functions on  $\Omega$ . Similarly  $H^s(\Omega, \mathbb{V})$  is the  $\mathbb{V}$ -valued  $H^s(\Omega)$  Sobolev space. Define the spaces

$$\begin{aligned} U &:= L^2(\Omega, \mathbb{R}^2) \text{ if } \Gamma_D \neq \emptyset, \quad U := L^2(\Omega, \mathbb{R}^2)/\mathcal{RM} \text{ if } \Gamma_D = \emptyset, \\ \Sigma(g) &:= \{\tau \in L^2(\Omega, \mathbb{S}) : \operatorname{div} \tau \in L^2(\Omega, \mathbb{R}^2), \tau n = g \text{ on } \Gamma_N\}. \end{aligned}$$

Given a subdomain  $\Omega_0 \subseteq \Omega$ , let  $(\cdot, \cdot)_{\Omega_0}$  denote the  $L^2(\Omega_0)$  inner product and  $(\cdot, \cdot) = (\cdot, \cdot)_{\Omega}$ . For a 1d submanifold  $\Gamma_0 \subseteq \bar{\Omega}$ , by  $\langle \cdot, \cdot \rangle_{\Gamma_0}$  we denote the  $L^2(\Gamma_0)$  inner product. The variational formulation of (1.2) seeks  $\sigma \in \Sigma(g_N)$  and  $u \in U$  such that

$$(\mathbb{A}\sigma, \tau) + (\operatorname{div} \tau, u) = \langle \tau n, g_D \rangle_{\Gamma_D}, \quad \tau \in \Sigma := \Sigma(0), \quad (1.3a)$$

$$(\operatorname{div} \sigma, v) = (f, v), \quad v \in U. \quad (1.3b)$$

Let  $\mathcal{T}_0$  be a conforming initial macro-triangulation of  $\Omega$  and be aligned with  $\Gamma_D, \Gamma_N$ . Let  $\mathbb{T} = \{\mathcal{T}_h\}$  denote a forest of conforming refinement of  $\mathcal{T}_0$  indexed by  $h$ . For  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$ , we say  $\mathcal{T}_H \leq \mathcal{T}_h$  provided  $\mathcal{T}_h$  is a refinement of  $\mathcal{T}_H$ . We assume  $\mathbb{T}$  is shape regular, *i.e.*, there exists a uniform constant  $\gamma_0$  such that

$$\max_{\mathcal{T}_h \in \mathbb{T}} \max_{T \in \mathcal{T}_h} r_T / \rho_T < \gamma_0 < \infty,$$

where  $r_T$  and  $\rho_T$  are radii of circumscribed and inscribed circles of  $T$ , respectively. Let  $\mathcal{P}_r(T, \mathbb{V})$  denote the space of  $\mathbb{V}$ -valued polynomials of degree at most  $r$  on  $T$ . For an integer  $r \geq 0$ , the mixed finite element spaces are

$$\begin{aligned} \Sigma_h(g) &:= \{\tau_h \in \Sigma(g) : \tau_h|_T \in \mathcal{P}_{r+3}(T, \mathbb{S}) \forall T \in \mathcal{T}_h\}, \\ U_h &:= \{v_h \in U : v_h|_T \in \mathcal{P}_{r+2}(T, \mathbb{R}^2) \forall T \in \mathcal{T}_h\}. \end{aligned}$$

In the sequel, we assume that  $g_N$  is a piecewise polynomial on  $\Gamma_N$  with  $g_N|_e \in \mathcal{P}_{r+3}(e, \mathbb{R}^2)$  for each edge  $e$  in  $\mathcal{T}_0$  and  $g_D$  is a piecewise polynomial on  $\Gamma_D$  aligned with  $\mathcal{T}_0$ . The mixed method for (1.3) is to find  $\sigma_h \in \Sigma_h(g_N), u_h \in U_h$  such that

$$(\mathbb{A}\sigma_h, \tau_h) + (\operatorname{div} \tau_h, u_h) = \langle \tau_h n, g_D \rangle_{\Gamma_D}, \quad \tau_h \in \Sigma_h := \Sigma_h(0), \quad (1.4a)$$

$$(\operatorname{div} \sigma_h, v) = (f, v_h), \quad v_h \in U_h, \quad (1.4b)$$

when  $\mathcal{T}_H \leq \mathcal{T}_h$ , it holds that  $\Sigma_H \times U_H \subseteq \Sigma_h \times U_h$ . Then using the nestedness and (1.4), we obtain the Galerkin orthogonality

$$(\mathbb{A}(\sigma_h - \sigma_H), \tau_H) + (\operatorname{div} \tau_H, u_h - u_H) = 0, \quad \tau_H \in \Sigma_H, \quad (1.5a)$$

$$(\operatorname{div}(\sigma_h - \sigma_H), v_H) = 0, \quad v_H \in U_H. \quad (1.5b)$$

It has been shown in [25, 30] that  $\Sigma_h \times U_h$  fulfills the inf-sup condition

$$\|v_h\| \leq C \sup_{0 \neq \tau_h \in \Sigma_h} \frac{(\operatorname{div} \tau_h, v_h)}{\|\tau_h\| + \|\operatorname{div} \tau_h\|}, \quad \forall v_h \in U_h, \quad (1.6)$$

where  $C$  depends only on  $r$  and  $\gamma_0$ . However, the construction of the local basis of  $\Sigma_h$  is rather involved [25]. To overcome this difficulty, (1.4) is implemented using hybridization technique and iterative solvers, see [25] and Section 6.

Let  $\mathcal{V}_h$  denote the set of grid vertices in  $\mathcal{T}_h$ . For  $r \geq 0$ , the classic Arnold–Winther mixed elasticity element spaces are

$$\Sigma_h^{\text{AW}} := \{\tau_h \in \Sigma : \tau_h|_T \in \mathcal{P}_{r+3}(T, \mathbb{S}), \operatorname{div} \tau \in \mathcal{P}_{r+1}(T, \mathbb{R}^2) \forall T \in \mathcal{T}_h, \tau_h \text{ is continuous at each } x \in \mathcal{V}_h\},$$

$$U_h^{\text{AW}} := \{v_h \in U : v_h|_T \in \mathcal{P}_{r+1}(T, \mathbb{R}^2) \ \forall T \in \mathcal{T}_h\}.$$

The Hu–Zhang mixed elasticity element spaces are

$$\begin{aligned} \Sigma_h^{\text{HZ}} &:= \{\tau_h \in \Sigma : \tau_h|_T \in \mathcal{P}_{r+3}(T, \mathbb{S}) \ \forall T \in \mathcal{T}_h, \tau_h \text{ is continuous at each } x \in \mathcal{V}_h\}, \\ U_h^{\text{HZ}} &:= U_h. \end{aligned}$$

Due to the continuity constraint of  $\Sigma_h^{\text{AW}}$  and  $\Sigma_h^{\text{HZ}}$  at each vertex, we note that  $\Sigma_H^{\text{AW}} \not\subseteq \Sigma_h^{\text{AW}}, \Sigma_H^{\text{HZ}} \not\subseteq \Sigma_h^{\text{HZ}}$ . This non-nestedness is the motivation of our analysis of adaptive hybridized MFEM and a major difficulty arising from the analysis of AMFEMs based on Arnold–Winther and Hu–Zhang elements.

The rest of this paper is organized as follows. In Section 2, we introduce preliminaries for deriving the discrete *a posteriori* upper error bound. In Section 3, we derive the discrete reliability and quasi-orthogonality. Section 4 is devoted to the convergence and optimality analysis of the proposed adaptive algorithm. In Section 5, we give proofs of technical results used in our analysis. The numerical experiment is presented in Section 6.

## 2. PRELIMINARIES

For  $T \in \mathcal{T}_h$ , let  $|T|$  denote the area of  $T$  and  $h_T = |T|^{\frac{1}{2}}$  the size of  $T$ . On  $\partial T$  let  $t$  be the counterclockwise unit tangent and  $n$  the outward unit normal to  $\partial T$ . On  $\partial\Omega$  let  $t$  be the counterclockwise unit tangent to  $\partial\Omega$ . In  $\mathcal{T}_h$ , let  $\mathcal{E}_h, \mathcal{E}_h^o, \mathcal{E}_h^D$  denote the set of edges, interior edges, and edges in  $\Gamma_D$ , respectively. Let

$$\mathcal{E}_h^o(T) = \{e \in \mathcal{E}_h^o : e \subset \partial T\}, \quad \mathcal{E}_h^D(T) = \{e \in \mathcal{E}_h^D : e \subset \partial T\}.$$

We use  $\|\cdot\|_{\Omega_0}$  to denote the  $L^2(\Omega_0)$  norm and  $\|\cdot\| = \|\cdot\|_{\Omega}$ . Each edge  $e$  in  $\mathcal{T}_h$  is assigned with a unit tangent  $t_e$  and a unit normal  $n_e$ . In addition,  $t_e$  is counterclockwise oriented and  $n_e$  is outward pointing provided  $e \subset \partial\Omega$ . If  $e$  is an interior edge shared by two triangles  $T_+$  and  $T_-$ , let  $[[\phi]]|_e = (\phi|_{T_+})|_e - (\phi|_{T_-})|_e$  denote the jump of  $\phi$  over  $e$ , where  $n_e$  is pointing from  $T_+$  to  $T_-$ .

Given a scalar-valued function  $w$  and a  $\mathbb{R}^2$ -valued function  $\phi = (\phi_1, \phi_2)$ , let

$$\text{curl } w := \left( -\frac{\partial w}{\partial x_2}, \frac{\partial w}{\partial x_1} \right)^\top, \quad \text{rot } \phi := \frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}.$$

For  $\mathbb{R}^2$ -valued  $v = (v_1, v_2)^\top$  and  $\mathbb{R}^{2 \times 2}$ -valued  $\tau = (\tau_1, \tau_2)^\top$ , let

$$\text{curl } v := (\text{curl } v_1, \text{curl } v_2)^\top, \quad \text{rot } \tau := (\text{rot } \tau_1, \text{rot } \tau_2)^\top.$$

For a unit vector  $d$ , we use  $\partial_d$  to denote the directional derivative along  $d$ . The stress error will be estimated by  $\eta_h = \eta_h(\sigma_h) = (\sum_{T \in \mathcal{T}_h} \eta_h^2(\sigma_h, T))^{\frac{1}{2}}$  with the element-wise error indicator given as

$$\begin{aligned} \eta_h(\sigma_h, T) = & \left\{ h_T^4 \|\text{rot rot } \mathbb{A}\sigma_h\|_T^2 + \sum_{e \in \mathcal{E}_h^o(T)} (h_e \|t_e^\top [[\mathbb{A}\sigma_h]] t_e\|_e^2 \right. \\ & + h_e^3 \|n_e^\top \partial_{t_e} [[\mathbb{A}\sigma_h]] t_e - [[\text{rot } \mathbb{A}\sigma_h] \cdot t_e\|_e^2) + \sum_{e \in \mathcal{E}_h^D(T)} (h_e \|t_e^\top ((\mathbb{A}\sigma_h)t_e - \partial_{t_e} g_D)\|_e^2 \\ & \left. + h_e^3 \|n_e^\top \partial_{t_e} (\mathbb{A}\sigma_h)t_e - (\text{rot } \mathbb{A}\sigma_h) \cdot t_e - n_e \cdot \partial_{t_e}^2 g_D\|_e^2) \right\}^{\frac{1}{2}}, \end{aligned}$$

where  $h_e$  is the diameter of  $e$ . By  $P_h$  we denote the  $L^2$  projection onto  $U_h$ . The data oscillation is  $\text{osc}_h = \text{osc}_h(f) = (\sum_{T \in \mathcal{T}_h} \text{osc}_h^2(f, T))^{\frac{1}{2}}$ , where

$$\text{osc}_h(f, T) = h_T \|f - P_h f\|_T.$$

The expression of  $\eta_h$  is the same as existing *a posteriori* error estimators for the MFEM using the Arnold–Winther and Hu–Zhang elements, see, e.g., [15, 20].

An indispensable ingredient of optimality analysis of AFEMs is the discrete upper bound for the finite element error. To construct such a bound, we consider the  $C^1$ -conforming space

$$W_h = \{w_h \in C^1(\bar{\Omega}) : w_h|_T \in \mathcal{P}_{r+5}(T) \ \forall T \in \mathcal{T}_h, (\text{curl } w_h)|_{\Gamma_j} \text{ is constant for } 1 \leq j \leq J\},$$

which is a subspace of the Morgan–Scott  $C^1$  element space [24, 37]. Let  $J$  denote the Airy stress function:

$$J = \text{curl curl} = \begin{pmatrix} \frac{\partial^2}{\partial x_2^2} & -\frac{\partial^2}{\partial x_1 \partial x_2} \\ -\frac{\partial^2}{\partial x_1 \partial x_2} & \frac{\partial^2}{\partial x_1^2} \end{pmatrix}.$$

Due to  $J(W_h) \subset \Sigma_h$  and  $\text{div} \circ J = 0$ , we obtain a well-defined discrete sequence:

$$W_h \xrightarrow{J} \Sigma_h \xrightarrow{\text{div}} U_h \longrightarrow 0. \tag{2.1}$$

**Lemma 2.1.** *The sequence (2.1) is exact, i.e.,  $\ker(\text{div}|_{\Sigma_h}) = J(W_h)$ .*

*Proof.* Given  $\tau_h \in \Sigma_h$  with  $\text{div } \tau_h = 0$ , there exists  $\phi \in H^1(\Omega)$  such that  $\text{curl } \phi = \tau_h$ . Due to the symmetry of  $\tau_h$ , it holds that  $\text{div } \phi = 0$  and thus  $\phi = \text{curl } w$  for some  $w \in H^2(\Omega)$ . Therefore we obtain  $\tau_h = \text{curl curl } w$ ,  $w|_T \in \mathcal{P}_{r+5}(T)$  for each  $T \in \mathcal{T}_h$ . The boundary condition  $\tau_h n|_{\Gamma_N} = 0$  implies  $\partial_t(\text{curl } w)|_{\Gamma_N} = 0$  and  $\text{curl } w$  is constant on each  $\Gamma_j$ . The proof is complete.  $\square$

The next theorem is a direct consequence of Lemma 2.1.

**Theorem 2.2** (Discrete Helmholtz decomposition).

$$\Sigma_h = J(W_h) \oplus \varepsilon_C^h(U_h),$$

where  $\varepsilon_C^h : U_h \rightarrow \Sigma_h$  is the adjoint operator of  $-\text{div} : \Sigma_h \rightarrow U_h$ , i.e.,

$$(\mathbb{A}\varepsilon_C^h(v_h), \tau_h) = -(v_h, \text{div } \tau_h) \text{ for all } \tau_h \in \Sigma_h.$$

*Proof.* Let  $\ker(\text{div}|_{\Sigma_h})^\perp$  be the orthogonal complement of  $\ker(\text{div}|_{\Sigma_h})$  in  $\Sigma_h$  with respect to the weighted inner product  $(\mathbb{A}\cdot, \cdot)$ . Elementary linear algebra shows that

$$\ker(\text{div}|_{\Sigma_h})^\perp = \varepsilon_C^h(U_h).$$

Combining it with the exactness  $\ker(\text{div}|_{\Sigma_h}) = J(W_h)$  in Lemma 2.1, we obtain

$$\Sigma_h = \ker(\text{div}|_{\Sigma_h}) \oplus \ker(\text{div}|_{\Sigma_h})^\perp = J(W_h) \oplus \varepsilon_C^h(U_h),$$

which completes the proof.  $\square$

**Remark 2.3.** For the Arnold–Winther and Hu–Zhang elements under  $\Gamma_N = \emptyset$ , the correct discrete elasticity sequences are

$$\widehat{W}_h \xrightarrow{J} \Sigma_h^{AW} \xrightarrow{\text{div}} U_h^{AW} \longrightarrow 0,$$

and

$$\widehat{W}_h \xrightarrow{J} \Sigma_h^{HZ} \xrightarrow{\text{div}} U_h \longrightarrow 0,$$

respectively, where

$$\widehat{W}_h = \{w_h \in H^2(\Omega) : w_h|_T \in \mathcal{P}_{r+5}(T) \text{ for each } T \in \mathcal{T}_h, \nabla^2 w_h \text{ is continuous at each } x \in \mathcal{V}_h\},$$

when  $r = 0$ ,  $\widehat{W}_h$  is the well-known quintic Argyris finite element space. Due to the extra vertex continuity, it is relatively easy to construct a local basis of  $\widehat{W}_h$  and interpolation onto  $\widehat{W}_h$ .

It is noted that  $W_h$  is not a standard finite element space. The work [37] gives a set of unisolvent nodal variables and locally supported dual nodal basis of  $W_h$ , although those degrees of freedom are much more complicated than the Argyris-type  $C^1$  space  $\widehat{W}_h$ . Based on a slightly modified (but complicated) nodal variables, Girault and Scott [24] constructed a locally defined and  $H^1$ -bounded interpolation preserving the homogeneous boundary condition. To derive the discrete reliability of (1.4), we present an interpolation  $I_H : W_h \rightarrow W_H$ , which is a slight variation of the interpolation in [24]. Throughout the rest of this paper, we say  $c_1 \lesssim c_2$  provided  $c_1 \leq cc_2$  for some generic constant  $c$  depending only on  $\mu, \Omega, \gamma_0, r$ .

**Lemma 2.4.** *For  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$ , let  $\mathcal{R}_H := \mathcal{T}_H \setminus \mathcal{T}_h$  be the set of refinement elements and*

$$\widetilde{\mathcal{R}}_H := \{T \in \mathcal{T}_H : T \cap T' \neq \emptyset \text{ for some } T' \in \mathcal{R}_H\}$$

*denote the enriched collection of refinement elements. There exists an interpolation  $I_H : W_h \rightarrow W_H$  such that for  $w_h \in W_h$ ,*

$$w_h - I_H w_h = 0 \quad \text{at } x \in \mathcal{V}_H, \tag{2.2a}$$

$$w_h - I_H w_h = 0 \quad \text{on } T \in \mathcal{T}_H \setminus \widetilde{\mathcal{R}}_H, \tag{2.2b}$$

$$w_h - I_H w_h = 0 \quad \text{on } \Gamma_N, \tag{2.2c}$$

$$\partial_n(w_h - I_H w_h) = 0 \quad \text{on } \Gamma_N. \tag{2.2d}$$

*In addition,*

$$\sum_{T \in \mathcal{T}_H} h_T^{-4} \|w_h - I_H w_h\|_T^2 + h_T^{-2} |w_h - I_H w_h|_{H^1(T)}^2 + h_T^{-3} \|w_h - I_H w_h\|_{\partial T}^2 + h_T^{-1} \|\nabla(w_h - I_H w_h)\|_{\partial T}^2 \lesssim |w_h|_{H^2(\Omega)}^2. \tag{2.3}$$

The proof of Proposition 2.4 is postponed in Section 5.

### 3. DISCRETE RELIABILITY AND QUASI-ORTHOGONALITY

Let  $\|\cdot\|_{\mathbb{A}}$  denote the norm corresponding to  $(\mathbb{A}, \cdot)$ . For  $\mathcal{M} \subseteq \mathcal{T}_h$ , let

$$\eta_h(\sigma_h, \mathcal{M}) = \left( \sum_{T \in \mathcal{M}} \eta_h^2(\sigma_h, T) \right)^{\frac{1}{2}},$$

$$\text{osc}_h(f, \mathcal{M}) = \left( \sum_{T \in \mathcal{M}} \text{osc}_h^2(f, T) \right)^{\frac{1}{2}}.$$

We shall prove the discrete reliability of the estimator  $\eta_h$  and quasi-orthogonality between  $\sigma - \sigma_h$  and  $\sigma_h - \sigma_H$ . The analysis relies on the following discrete approximation result, whose proof is left in Section 5.

**Lemma 3.1.** *Let  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$  and  $Q_H$  denote the  $L^2$ -projection onto the space of piecewise rigid body motions*

$$\mathcal{RM}_H = \{v \in L^2(\Omega, \mathbb{R}^2) : v|_T \in \mathcal{RM} \text{ for all } T \in \mathcal{T}_H\}.$$

*It holds that*

$$\left( \sum_{T \in \mathcal{T}_H} h_T^{-2} \|v_h - Q_H v_h\|_T^2 \right)^{\frac{1}{2}} \lesssim \|\varepsilon_C^h(v_h)\|_{\mathbb{A}}.$$

The space  $\mathcal{RM}_H$  can be viewed as a broken rotated Raviart–Thomas finite element space. Here we are interested in  $Q_H$  instead of  $P_H$  because we will use the fact  $\mathcal{RM} \subset \ker(\varepsilon)$ , see the proof of Lemma 5.2 for details.

The next lemma is used to remove the Lamé coefficient  $\lambda$  in error bounds. The proof can be found in Lemmas 3.1 and 3.2 of [3].

**Lemma 3.2.** *There exists a constant  $C_{\text{rb}}$  depends only on  $\mu$  and  $\Omega$ , such that*

$$\|\tau\| \leq C_{\text{rb}}(\|\tau\|_{\mathbb{A}} + \|\operatorname{div} \tau\|_{H^{-1}(\Omega)})$$

for all  $\tau \in \Sigma$  with  $\int_{\Omega} \operatorname{tr} \tau dx = 0$ .

For  $w \in H^1(T)$  and  $\phi \in H^1(T, \mathbb{R}^2)$ , we have the integration-by-parts formula:

$$(\operatorname{curl} w, \phi)_T = \langle w, \phi \cdot t \rangle_{\partial T} - (w, \operatorname{rot} \phi)_T. \tag{3.1}$$

With the above preparations, we are able to prove the discrete reliability of  $\eta_h$ .

**Theorem 3.3** (Discrete reliability). *Let  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$ . There exists a constant  $C_{\text{drel}}$  depending only on  $\mu, \Omega, \gamma_0$ , such that*

$$\|\sigma_H - \sigma_h\|_{\mathbb{A}}^2 \leq C_{\text{drel}} \left( \eta_H^2(\sigma_H, \tilde{\mathcal{R}}_H) + \operatorname{osc}_H^2(f, \mathcal{R}_H) \right).$$

*Proof.* Applying Theorem 2.2 to  $\sigma_H - \sigma_h \in \Sigma_H$  gives

$$\sigma_H - \sigma_h = J(w_h) + \varepsilon_{\mathbb{C}}^h(v_h) \tag{3.2}$$

for some  $w_h \in W_h$  and  $v_h \in U_h$ . Taking  $\tau_H = \delta$  in (1.5a) leads to

$$\int_{\Omega} \operatorname{tr}(\sigma_H - \sigma_h) dx = 0. \tag{3.3}$$

Direct calculation shows that

$$\int_{\Omega} \frac{1}{2(\mu + \lambda)} \operatorname{tr} \varepsilon_{\mathbb{C}}^h(v_h) dx = (\mathbb{A} \varepsilon_{\mathbb{C}}^h(v_h), \delta) = -(v_h, \operatorname{div} \delta) = 0. \tag{3.4}$$

Then a combination of (3.2)–(3.4) yields

$$\int_{\Omega} \operatorname{tr} J(w_h) dx = 0. \tag{3.5}$$

Hence using Lemma 3.2,  $\operatorname{div} \circ J = 0$ , (3.5) and the  $\mathbb{A}$ -orthogonality between  $Jw_h$  and  $\varepsilon_{\mathbb{C}}^h(v_h)$ , we obtain the following robust bound

$$\|Jw_h\| \lesssim \|Jw_h\|_{\mathbb{A}} \leq \|\sigma_H - \sigma_h\|_{\mathbb{A}}. \tag{3.6}$$

Let  $E_h = w_h - I_H w_h$ . Using (1.4a), (1.5a),  $\operatorname{div} \circ J = 0$ , (2.2b), we have

$$\begin{aligned} (\mathbb{A}(\sigma_H - \sigma_h), Jw_h) &= (\mathbb{A}(\sigma_H - \sigma_h), JE_h) \\ &= (\mathbb{A}\sigma_H, JE_h) + \langle \partial_t \operatorname{curl} E_h, g_D \rangle_{\Gamma_D} \\ &= \sum_{T \in \tilde{\mathcal{R}}_H} (\mathbb{A}\sigma_H, JE_h)_T - \langle \operatorname{curl} E_h, \partial_t g_D \rangle_{\Gamma_D}. \end{aligned} \tag{3.7}$$

In the last equality, we integrate by parts on  $\Gamma_D$  and use the fact  $\operatorname{curl} E_h = 0$  on  $\partial\Gamma_D \subset \bar{\Gamma}_N$  (see (2.2c), (2.2d)). For each  $T \in \tilde{\mathcal{R}}_H$ , using the formula (3.1), we have

$$\begin{aligned} (\mathbb{A}\sigma_H, JE_h)_T &= \langle (\mathbb{A}\sigma_H)t, \operatorname{curl} E_h \rangle_{\partial T} - (\operatorname{rot} \mathbb{A}\sigma_H, \operatorname{curl} E_h)_T \\ &= \langle (\mathbb{A}\sigma_H)t, \operatorname{curl} E_h \rangle_{\partial T} - \langle (\operatorname{rot} \mathbb{A}\sigma_H) \cdot t, E_h \rangle_{\partial T} + (\operatorname{rot} \operatorname{rot} \mathbb{A}\sigma_H, E_h)_T. \end{aligned} \tag{3.8}$$

Integrating by parts on each edge of  $T$ , we have

$$\begin{aligned} \langle (\mathbb{A}\sigma_H)t, \text{curl } E_h \rangle_{\partial T} &= \langle t^\top (\mathbb{A}\sigma_H)t, \partial_n E_h \rangle_{\partial T} - \langle n^\top (\mathbb{A}\sigma_H)t, \partial_t E_h \rangle_{\partial T} \\ &= \langle t^\top (\mathbb{A}\sigma_H)t, \partial_n E_h \rangle_{\partial T} + \langle n^\top \partial_t (\mathbb{A}\sigma_H)t, E_h \rangle_{\partial T}. \end{aligned} \quad (3.9)$$

In the last equality, we use the property (2.2a), *i.e.*,  $E_h = 0$  at each vertex of  $T$ . Similarly, for the boundary term in (3.7),

$$\begin{aligned} -\langle \text{curl } E_h, \partial_t g_D \rangle_{\Gamma_D} &= -\langle \partial_n E_h, t \cdot \partial_t g_D \rangle_{\Gamma_D} + \langle \partial_t E_h, n \cdot \partial_t g_D \rangle_{\Gamma_D} \\ &= -\langle \partial_n E_h, t \cdot \partial_t g_D \rangle_{\Gamma_D} - \langle E_h, n \cdot \partial_t^2 g_D \rangle_{\Gamma_D}. \end{aligned} \quad (3.10)$$

Let  $\mathcal{E}_H^o(\tilde{\mathcal{R}}_H) = \{e \in \mathcal{E}_H^o : e \subset \partial T \text{ for some } T \in \tilde{\mathcal{R}}_H\}$  and  $\mathcal{E}_H^D(\tilde{\mathcal{R}}_H) = \{e \in \mathcal{E}_H : e \subset \partial T \text{ for some } T \in \tilde{\mathcal{R}}_H, e \subset \Gamma_D\}$ . Note that  $E_h$  and  $\partial_n E_h$  are continuous over each edge in  $\mathcal{T}_H$ . Combining (3.8)–(3.10) and using (2.2b), (2.2c), (2.2d), we obtain

$$\begin{aligned} (\mathbb{A}(\sigma_H - \sigma_h), Jw_h) &= \sum_{T \in \tilde{\mathcal{R}}_H} \left\{ \langle t^\top (\mathbb{A}\sigma_H)t, \partial_n E_h \rangle_{\partial T} \right. \\ &\quad \left. + (\text{rot rot } \mathbb{A}\sigma_H, E_h)_T + \langle n^\top \partial_t (\mathbb{A}\sigma_H)t - (\text{rot } \mathbb{A}\sigma_H) \cdot t, E_h \rangle_{\partial T} \right\} \\ &\quad - \sum_{e \in \mathcal{E}_H^D(\tilde{\mathcal{R}}_H)} \left\{ \langle \partial_{n_e} E_h, t_e \cdot \partial_{t_e} g_D \rangle_e + \langle E_h, n_e \cdot \partial_{t_e}^2 g_D \rangle_e \right\} \\ &= \sum_{e \in \mathcal{E}_H^o(\tilde{\mathcal{R}}_H)} \left\{ \langle t_e^\top [\mathbb{A}\sigma_H]t_e, \partial_{n_e} E_h \rangle_e + \langle n_e^\top \partial_{t_e} [\mathbb{A}\sigma_H]t_e - [\text{rot } \mathbb{A}\sigma_H] \cdot t_e, E_h \rangle_e \right\} \\ &\quad + \sum_{T \in \tilde{\mathcal{R}}_H} (\text{rot rot } \mathbb{A}\sigma_H, E_h)_T + \sum_{e \in \mathcal{E}_H^D(\tilde{\mathcal{R}}_H)} \left\{ \langle t_e^\top (\mathbb{A}\sigma_H)t_e - \partial_{t_e} g_D, \partial_{n_e} E_h \rangle_e \right. \\ &\quad \left. + \langle n_e^\top \partial_{t_e} (\mathbb{A}\sigma_H)t_e - (\text{rot } \mathbb{A}\sigma_H) \cdot t_e - n_e^\top \partial_{t_e}^2 g_D, E_h \rangle_e \right\}. \end{aligned} \quad (3.11)$$

Using the expression (3.11) and the Cauchy–Schwarz inequality, we have

$$(\mathbb{A}(\sigma_H - \sigma_h), Jw_h) \lesssim \eta_H(\sigma_H, \tilde{\mathcal{R}}_H) \left( \sum_{T \in \mathcal{T}_H} h_T^{-4} \|E_h\|_T^2 + h_T^{-1} \|\partial_n E_h\|_{\partial T}^2 + h_T^{-3} \|E_h\|_{\partial T}^2 \right)^{\frac{1}{2}}. \quad (3.12)$$

It then follows from (3.12), (2.3) and (3.6) that

$$\begin{aligned} (\mathbb{A}(\sigma_H - \sigma_h), Jw_h) &\lesssim \eta_H(\sigma_H, \tilde{\mathcal{R}}_H) |w_h|_{H^2(\Omega)} \\ &\lesssim \eta_H(\sigma_H, \tilde{\mathcal{R}}_H) \|\sigma_h - \sigma_H\|_{\mathbb{A}}. \end{aligned} \quad (3.13)$$

On the other hand, equation (1.4b) implies

$$\begin{aligned} (\mathbb{A}(\sigma_H - \sigma_h), \varepsilon_C^h(v_h)) &= -(\text{div}(\sigma_H - \sigma_h), v_h) = (P_h f - P_H f, v_h) \\ &= (f - P_H f, v_h - Q_H v_h) = \sum_{T \in \mathcal{R}_H} (f - P_H f, v_h - Q_H v_h)_T. \end{aligned} \quad (3.14)$$

Using (3.14), Lemma 3.1, and  $\|\varepsilon_C^h(v_h)\|_{\mathbb{A}} \leq \|\sigma_H - \sigma_h\|_{\mathbb{A}}$ , we obtain

$$\begin{aligned} (\mathbb{A}(\sigma_H - \sigma_h), \varepsilon_C^h(v_h)) &\leq \text{osc}_H(f, \mathcal{R}_H) \left( \sum_{T \in \mathcal{R}_H} h_T^{-2} \|v_h - Q_H v_h\|_T^2 \right)^{\frac{1}{2}} \\ &\lesssim \text{osc}_H(f, \mathcal{R}_H) \|\varepsilon_C^h(v_h)\|_{\mathbb{A}} \leq \text{osc}_H(f, \mathcal{R}_H) \|\sigma_H - \sigma_h\|_{\mathbb{A}}. \end{aligned} \quad (3.15)$$

Finally, a combination of (3.13) and (3.15) completes the proof.  $\square$



Let  $\mathcal{T}_h$  be a uniform refinement of  $\mathcal{T}_H$  and let the maximum mesh size of  $\mathcal{T}_h$  go to 0 in Theorem 3.3. In this case,  $\tilde{\mathcal{R}}_H = \mathcal{R}_H = \mathcal{T}_H$  and  $\sigma_h \rightarrow \sigma, u_h \rightarrow u$  in  $\Sigma \times U$ . Therefore, we obtain the continuous upper bound

$$\|\sigma - \sigma_H\|_{\mathbb{A}}^2 \leq C_{\text{rel}}(\eta_H^2(\sigma_H) + \text{osc}_H^2(f)), \tag{3.16}$$

where  $C_{\text{rel}} \in (0, C_{\text{drel}}]$  is a constant depending only on  $\mu, \Omega, \gamma_0$ .

The quasi-orthogonality on the variable  $\sigma$  follows with a similar argument as in the proof of Theorem 3.3.

**Theorem 3.4** (Quasi-orthogonality). *Let  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$ . For any  $\nu \in (0, 1)$ , it holds that*

$$(1 - \nu)\|\sigma - \sigma_h\|_{\mathbb{A}}^2 \leq \|\sigma - \sigma_H\|_{\mathbb{A}}^2 - \|\sigma_h - \sigma_H\|_{\mathbb{A}}^2 + C_\nu \text{osc}_H^2(f, \mathcal{R}_H),$$

where  $C_\nu = \nu^{-1}C_\sigma$  and  $C_\sigma$  is a constant depending only on  $\mu, \Omega, \gamma_0$ .

*Proof.* Combining (3.2) and (1.3a), (1.4a) yields

$$(\mathbb{A}(\sigma - \sigma_h), \sigma_H - \sigma_h) = (\mathbb{A}(\sigma - \sigma_h), \varepsilon_C^h(v_h)) \leq \|\sigma - \sigma_h\|_{\mathbb{A}} \|\varepsilon_C^h(v_h)\|_{\mathbb{A}}. \tag{3.17}$$

Following the same analysis in (3.14), we have

$$\begin{aligned} \|\varepsilon_C^h(v_h)\|_{\mathbb{A}}^2 &= -(\text{div } \varepsilon_C^h(v_h), v_h) = -(\text{div}(\sigma_H - \sigma_h), v_h) \\ &= \sum_{T \in \mathcal{R}_H} (f - P_H f, v_h - Q_H v_h)_T \leq C_\sigma^{\frac{1}{2}} \text{osc}_H(f, \mathcal{R}_H) \|\varepsilon_C^h(v_h)\|_{\mathbb{A}}. \end{aligned} \tag{3.18}$$

A combination of (3.17) and (3.18) shows that

$$\begin{aligned} (\mathbb{A}(\sigma - \sigma_h), \sigma_H - \sigma_h) &\leq C_\sigma^{\frac{1}{2}} \|\sigma - \sigma_h\|_{\mathbb{A}} \text{osc}_H(f, \mathcal{R}_H) \\ &\leq \frac{\nu}{2} \|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \frac{\nu^{-1}}{2} C_\sigma \text{osc}_H^2(f, \mathcal{R}_H), \end{aligned}$$

where  $0 < \nu < 1$ . Therefore

$$\begin{aligned} \|\sigma - \sigma_h\|_{\mathbb{A}}^2 &= \|\sigma - \sigma_H\|_{\mathbb{A}}^2 - \|\sigma_h - \sigma_H\|_{\mathbb{A}}^2 + 2(\mathbb{A}(\sigma - \sigma_h), \sigma_H - \sigma_h) \\ &\leq \|\sigma - \sigma_H\|_{\mathbb{A}}^2 - \|\sigma_h - \sigma_H\|_{\mathbb{A}}^2 + \nu \|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \nu^{-1} C_\sigma \text{osc}_H^2(f, \mathcal{R}_H). \end{aligned}$$

The proof is complete. □

#### 4. QUASI-OPTIMALITY

Define

$$\begin{aligned} \bar{\eta}_h(\sigma_h, T) &= (\eta_h^2(\sigma_h, T) + \text{osc}_h^2(f, T))^{\frac{1}{2}}, \\ \bar{\eta}_h(\sigma_h, \mathcal{M}) &= \left( \sum_{T \in \mathcal{M}} \bar{\eta}_h^2(\sigma_h, T) \right)^{\frac{1}{2}}, \quad \mathcal{M} \subseteq \mathcal{T}_h. \end{aligned}$$

Our adaptive algorithm is based on the classical feedback loop

$$\text{SOLVE} \longrightarrow \text{ESTIMATE} \longrightarrow \text{MARK} \longrightarrow \text{REFINE}.$$

In the procedure REFINE, we use the newest vertex bisection [5, 36, 40] to ensure the shape regularity of  $\mathbb{T}$ .

**Algorithm 4.1.** Input the initial mesh  $\mathcal{T}_{h_0} = \mathcal{T}_0$  and  $\theta \in (0, 1)$ . Set  $\ell = 0$ .

**SOLVE:** Solve (1.4) on  $\mathcal{T}_{h_\ell}$  to obtain the finite element solution  $(\sigma_{h_\ell}, u_{h_\ell})$ .

**ESTIMATE:** Calculate error indicators  $\{\bar{\eta}_{h_\ell}(\sigma_{h_\ell}, T)\}_{T \in \mathcal{T}_{h_\ell}}$ .

**MARK:** Select a subset  $\mathcal{M}_\ell \subset \mathcal{T}_{h_\ell}$  with minimal cardinality such that

$$\bar{\eta}_{h_\ell}(\sigma_{h_\ell}, \mathcal{M}_\ell) \geq \theta \bar{\eta}_{h_\ell}.$$

**REFINE:** Refine all elements in  $\mathcal{M}_\ell$  and minimal number of neighboring elements to remove hanging nodes. The resulting conforming mesh is  $\mathcal{T}_{h_{\ell+1}}$ . Set  $\ell = \ell + 1$ . Go to **SOLVE**.

In the procedure **ESTIMATE**, the actual estimator is  $(\eta_{h_\ell}^2 + \text{osc}_{h_\ell}^2)^{\frac{1}{2}}$  instead of  $\eta_{h_\ell}$ . Due to this strategy, an extra marking step for data oscillation can be avoided, see, e.g., [27, 31]. Since the data oscillation  $\text{osc}_{h_\ell}$  is completely local, its behavior can be easily described by the following lemma, see, e.g., Lemma 5.2 in [27].

**Lemma 4.2.** *For  $\ell \geq 0$ , let  $\mathcal{R}_\ell = \mathcal{T}_{h_\ell} \setminus \mathcal{T}_{h_{\ell+1}}$  denote the collection of refinement elements from  $\mathcal{T}_{h_\ell}$  to  $\mathcal{T}_{h_{\ell+1}}$ . It holds that*

$$\text{osc}_{h_{\ell+1}}^2 \leq \text{osc}_{h_\ell}^2 - \frac{1}{2} \text{osc}_{h_\ell}^2(f, \mathcal{R}_\ell).$$

Once the theoretical results in Section 3 are available, the convergence and complexity analysis of Algorithm 4.1 follows from classical and systematic arguments, see, e.g., [14, 16, 39] for axioms of adaptivity. To be self-contained, we still briefly outlined the rest of adaptivity analysis.

The estimator reduction is a standard ingredient in the convergence analysis of AFEMs, see [16]. Since  $\bar{\eta}_h$  involves data oscillation, we refer to Lemma 5.1 in [27] for a detailed proof. The only new ingredient in the proof of Lemma 4.3 is the robust inequality  $\|\mathbb{A}\tau\| \lesssim \|\tau\|_{\mathbb{A}}$ .

**Lemma 4.3.** *There exists a constant  $\gamma \in (0, 1)$  and  $C_{\text{re}} > 0$  depending only on  $\mu, \Omega, \gamma_0, \theta$  such that*

$$\bar{\eta}_{h_{\ell+1}}^2 \leq \gamma \bar{\eta}_{h_\ell}^2 + C_{\text{re}} \|\sigma_{h_\ell} - \sigma_{h_{\ell+1}}\|_{\mathbb{A}}^2.$$

For convenience, let

$$e_\ell = \|\sigma - \sigma_{h_\ell}\|_{\mathbb{A}}, \quad E_\ell = \|\sigma_{h_\ell} - \sigma_{h_{\ell+1}}\|_{\mathbb{A}}.$$

The next theorem gives the contraction property of Algorithm 4.1.

**Theorem 4.4 (Contraction).** *There exists constants  $\nu, \alpha \in (0, 1)$  depending only on  $\theta, \mu, \Omega, \gamma_0$  such that*

$$(1 - \nu)e_{\ell+1}^2 + 2C_\nu \text{osc}_{h_{\ell+1}}^2 + C_{\text{re}}^{-1} \bar{\eta}_{h_{\ell+1}}^2 \leq \alpha((1 - \nu)e_\ell^2 + 2C_\nu \text{osc}_{h_\ell}^2 + C_{\text{re}}^{-1} \bar{\eta}_{h_\ell}^2).$$

*Proof.* A combination of Theorem 3.4 and Lemma 4.2 shows that

$$(1 - \nu)e_{\ell+1}^2 + 2C_\nu \text{osc}_{h_{\ell+1}}^2 \leq e_\ell^2 + 2C_\nu \text{osc}_{h_\ell}^2 - E_{h_\ell}^2. \quad (4.1)$$

On the other hand, the reliability (3.16) gives

$$e_\ell^2 \leq C_{\text{rel}} \bar{\eta}_{h_\ell}^2, \quad \text{osc}_{h_\ell}^2 \leq \bar{\eta}_{h_\ell}^2. \quad (4.2)$$

Let  $\alpha \in (0, 1)$  be a constant. Using (4.1) and Lemma 4.3, we have

$$\begin{aligned} (1 - \nu)e_{\ell+1}^2 + 2C_\nu \text{osc}_{h_{\ell+1}}^2 + C_{\text{re}}^{-1} \bar{\eta}_{h_{\ell+1}}^2 &\leq e_\ell^2 + 2C_\nu \text{osc}_{h_\ell}^2 + C_{\text{re}}^{-1} \gamma \bar{\eta}_{h_\ell}^2 \\ &\leq \alpha(1 - \nu)e_\ell^2 + 2\alpha C_\nu \text{osc}_{h_\ell}^2 + (1 - \alpha(1 - \nu))e_\ell^2 + 2(1 - \alpha)C_\nu \text{osc}_{h_\ell}^2 + C_{\text{re}}^{-1} \gamma \bar{\eta}_{h_\ell}^2. \end{aligned}$$

Combining it with (4.2) yields

$$(1 - \nu)e_{\ell+1}^2 + 2C_\nu \text{osc}_{h_{\ell+1}}^2 + C_{\text{re}}^{-1} \bar{\eta}_{h_{\ell+1}}^2 \leq \alpha(1 - \nu)e_\ell^2 + 2\alpha C_\nu \text{osc}_{h_\ell}^2 + \{(1 - \alpha(1 - \nu))C_{\text{rel}} + 2(1 - \alpha)C_\nu + C_{\text{re}}^{-1}\gamma\} \bar{\eta}_{h_\ell}^2. \tag{4.3}$$

Let  $(1 - \alpha(1 - \nu))C_{\text{rel}} + 2(1 - \alpha)C_\nu + C_{\text{re}}^{-1}\gamma = \alpha C_{\text{re}}^{-1}$ , i.e.,

$$\alpha = \frac{C_{\text{rel}} + 2C_\nu + C_{\text{re}}^{-1}\gamma}{(1 - \nu)C_{\text{rel}} + 2C_\nu + C_{\text{re}}^{-1}}.$$

Clearly  $\alpha < 1$  provided  $0 < \nu < \frac{1-\gamma}{C_{\text{re}}C_{\text{rel}}}$ . The contraction follows from (4.3). □

The efficiency of  $\bar{\eta}_h$  follows with the same bubble function technique in [15, 20]:

$$C_{\text{eff}} \bar{\eta}_h^2(\sigma_h) \leq \|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \text{osc}_h^2(f), \tag{4.4}$$

where the constant  $C_{\text{eff}} > 0$  depends only on  $\mu, \Omega, \gamma_0, r$ . An essential ingredient in the complexity analysis is the following cardinality estimate

$$\#\mathcal{T}_{h_\ell} - \#\mathcal{T}_{h_0} \lesssim \sum_{j=0}^{\ell-1} \mathcal{M}_j. \tag{4.5}$$

It has been shown in [7] that (4.5) holds provided the newest vertices in the initial mesh  $\mathcal{T}_{h_0}$  are suitably chosen. In addition, the marking parameter  $\theta$  is required to be below the threshold

$$\theta_* = \min\left(1, \frac{C_{\text{eff}}}{3C_{\text{drel}}}\right)^{\frac{1}{2}},$$

which can be derived in the next lemma.

**Lemma 4.5** (Optimal marking). *Let  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$ . Set  $\mu = \frac{1}{2}\left(1 - \frac{\theta^2}{\theta_*^2}\right)$ . If*

$$\|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \text{osc}_h^2(f) \leq \mu\{\|\sigma - \sigma_H\|_{\mathbb{A}}^2 + \text{osc}_H^2(f)\}. \tag{4.6}$$

*Then the set  $\tilde{\mathcal{R}}_H$  in Lemma 2.4 verifies the Dörfler marking property*

$$\bar{\eta}_H(\sigma_H, \tilde{\mathcal{R}}_H) \geq \theta \bar{\eta}_H.$$

*Proof.* Using (4.4) and (4.6), we have

$$\begin{aligned} (1 - 2\mu)C_{\text{eff}}\bar{\eta}_H^2 &\leq (1 - 2\mu)(\|\sigma - \sigma_H\|_{\mathbb{A}}^2 + \text{osc}_H^2(f)) \\ &\leq \|\sigma - \sigma_H\|_{\mathbb{A}}^2 - 2\|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \text{osc}_H^2(f) - 2\text{osc}_h^2(f) \\ &\leq 2\|\sigma_H - \sigma_h\|_{\mathbb{A}}^2 + \text{osc}_H^2(f, \mathcal{R}_H). \end{aligned} \tag{4.7}$$

In the last step, we use the obvious inequality

$$\text{osc}_H^2(f) - 2\text{osc}_h^2(f) \leq \text{osc}_H^2(f) - \text{osc}_h^2(f) \leq \text{osc}_H^2(f, \mathcal{R}_H).$$

It then follows from (4.7) and Theorem 3.3 that

$$(1 - 2\mu)C_{\text{eff}}\bar{\eta}_H^2 \leq 3C_{\text{drel}}\bar{\eta}_H^2(\sigma_H, \tilde{\mathcal{R}}_H).$$

The proof is then complete by  $\theta_*^2 \leq \frac{C_{\text{eff}}}{3C_{\text{drel}}}$ . □

Under these assumptions, the convergence rate of can be characterized by the nonlinear approximation property of  $\sigma$  and  $f$ . Let  $\tilde{\mathbb{T}}$  be the collection of grids created by newest vertex bisection from  $\mathcal{T}_{h_0}$ . For  $s > 0$ , define the semi-norms

$$|\sigma|_s = \sup_{N>0} \left\{ N^s \min_{\mathcal{T}_h \in \tilde{\mathbb{T}}, \#\mathcal{T}_h - \#\mathcal{T}_{h_0} \leq N} \min_{\tau_h \in \Sigma_h} \|\sigma - \tau_h\|_{\mathbb{A}} \right\},$$

$$|f|_s^o = \sup_{N>0} \left\{ N^s \min_{\mathcal{T}_h \in \tilde{\mathbb{T}}, \#\mathcal{T}_h - \#\mathcal{T}_{h_0} \leq N} \text{osc}_h(f) \right\}.$$

One can also define the coupled approximation semi-norm

$$|(\sigma, f)|_s := \sup_{N>0} \left\{ N^s \min_{\mathcal{T}_h \in \tilde{\mathbb{T}}, \#\mathcal{T}_h - \#\mathcal{T}_{h_0} \leq N} \left( \|\sigma - \sigma_h\|_{\mathbb{A}}^2 + \text{osc}_h(f)^2 \right)^{\frac{1}{2}} \right\}.$$

Since  $\lambda, \mu$  are constants, we have the following equivalence

$$|(\sigma, f)|_s < \infty \Leftrightarrow |\sigma|_s + |f|_s^o < \infty$$

as argued in Lemma 5.3 of [16]. The quasi-optimal convergence rate of Algorithm 4.1 follows from the contraction and previous assumptions, see, *e.g.*, [16], Lemma 5.10 and Theorems 5.1 for details.

**Theorem 4.6** (Quasi-optimality). *Let  $\{(\sigma_{h_\ell}, u_{h_\ell}, \mathcal{T}_{h_\ell})\}_{\ell \geq 0}$  be a sequence of finite element solutions and meshes generated by Algorithm 4.1. Assume  $|\sigma|_s + |f|_s^o < \infty$ ,  $\theta \in (0, \theta_*)$ , and (4.5) hold. There exists a constant  $C_{\text{opt}}$  depending only on  $\theta, \theta_*, \alpha, \mu, \Omega, \gamma_0$ , such that*

$$\left( \|\sigma - \sigma_{h_\ell}\|_{\mathbb{A}}^2 + \text{osc}_{h_\ell}^2 \right)^{\frac{1}{2}} \leq C_{\text{opt}} (|\sigma|_s + |f|_s^o) (\#\mathcal{T}_{h_\ell} - \#\mathcal{T}_{h_0})^{-s}.$$

## 5. LOCAL INTERPOLATION AND DISCRETE APPROXIMATION

In this section, we give proofs of Proposition 2.4 and Lemma 3.1. The  $H^1$ -bounded regularized interpolation in [24] does not satisfy the property (2.2a). For our purpose, an  $H^2$ -bounded interpolation is enough and we will not regularize the degrees of freedom based on point evaluation.

*Proof of Proposition 2.4.* Given  $x \in \mathcal{V}_H$ ,  $e \in \mathcal{E}_H$ ,  $w_H \in W_H$ , we say  $\partial_e \partial_e w_H(x) = \partial_e^2 w_H(x)$  is a second edge derivative at  $x$ , where  $\partial_e = \partial_{t_e}$  is the directional derivative along  $t_e$ . For two edges  $e, e'$  of  $T \in \mathcal{T}_H$  having  $x$ ,  $\partial_e \partial_{e'}(w_H|_T)(x)$  is called a cross derivative at  $x$ . A vertex  $x$  is called singular if all edges in  $\mathcal{E}_H$  meeting at  $x$  fall on two straight lines. The nodal variables (global degrees of freedom) of  $w \in W_H$  given in [24] are briefly described as follows.

- (1) the value of  $w$  and  $\nabla w$  at each vertex  $x \in \mathcal{V}_H$ ;
- (2) the edge normal derivative  $\partial_{n_e} w$  at  $r+1$  distinct interior points of each edge  $e \in \mathcal{E}_H$ ;
- (3) the value of  $w$  at  $r$  distinct interior points of each edge  $e \in \mathcal{E}_H$ ;
- (4) the value of  $w$  at  $r(r+1)/2$  distinct interior points of each triangle  $T \in \mathcal{T}_H$ ;
- (5) one cross derivative of  $w$  at each vertex  $x \in \mathcal{V}_H$  and two cross derivatives of  $w$  at each nonsingular boundary vertex  $x \in \mathcal{V}_H$ ;
- (6) the second edge derivatives of  $w$  for all edges meeting at each vertex  $x \in \mathcal{V}_H$ , with the exception of one interior edge per nonsingular vertex.

To preserve boundary conditions, for each boundary vertex, one edge used for defining edge and cross derivatives at that vertex is chosen to be on  $\partial\Omega$ .

Let  $\{w \rightarrow D_i w(a_i)\}_{i=1}^N$  denote the collection of the above nodal variables, where a node  $a_i$  could be a vertex in  $\mathcal{T}_H$  or an interior point of an edge/triangle in  $\mathcal{T}_H$ ,  $D_i$  is a differential operator of order  $|D_i| = 0$  (point evaluation), 1 (edge normal derivative), or 2 (second edge derivative and cross derivative). Note that  $\{a_i\}_{i=1}^N$  are not distinct since a node may be associated with multiple differential operators.

If  $a_i \in \mathcal{V}_H$  and  $D_i = \partial_{e_i^1} \partial_{e_i^2}$  is a second edge derivative or cross derivative, let  $e_i = e_i^1 \ni a_i$ . If  $a_i \in \mathcal{V}_H$  and  $D_i = \partial_{x_1}$  or  $\partial_{x_2}$ , let  $e_i \ni a_i$  be any edge in  $\mathcal{E}_H$ . If  $a_i$  is an interior point of  $e \in \mathcal{E}_H$ , we choose  $e_i = e$ . Moreover,  $e_i$  is chosen to be in  $\partial\Omega$  if  $a_i$  is a boundary vertex. By Riesz's representation theorem, there exists a polynomial  $\psi_i \in \mathcal{P}_{r+5}(e_i)$ , such that

$$\int_{e_i} w \psi_i b_i ds = w(a_i) \text{ for all } w \in \mathcal{P}_{r+5}(e_i), \tag{5.1}$$

where  $b_i$  is the edge bubble polynomial of unit size vanishing on the boundary of  $e_i$ . For a second edge derivative or cross derivative  $\partial_{e_i^1} \partial_{e_i^2}$ , using (5.1) and integrating by parts formula on  $e_i$ , we have

$$\begin{aligned} (\partial_{e_i^1} \partial_{e_i^2} w)(a_i) &= \int_{e_i} \partial_{e_i^1} \partial_{e_i^2} w(a_i) \psi_i b_i ds \\ &= - \int_{e_i} \partial_{e_i^2} w(a_i) \partial_{e_i^1} (\psi_i b_i) ds \text{ for all } w \in \mathcal{P}_{r+5}(e_i). \end{aligned} \tag{5.2}$$

Let  $\{\phi_i\}_{i=1}^N$  be the basis dual to the unisolvent set  $\{w \rightarrow D_i w(a_i)\}_{i=1}^N$ . For  $w_h \in W_h$ , we define the interpolant  $I_H w_h$  as

$$\begin{aligned} I_H w_h &= \sum_{|D_i|=0} w_h(a_i) \phi_i + \sum_{|D_i|=1} \left( \int_{e_i} (D_i w_h) \psi_i b_i ds \right) \phi_i \\ &\quad - \sum_{|D_i|=2} \left( \int_{e_i} \partial_{e_i^2} w_h(a_i) \partial_{e_i^1} (\psi_i b_i) ds \right) \phi_i. \end{aligned} \tag{5.3}$$

The definition of  $I_H$  implies  $w_h - I_H w_h = 0$  at each node  $a_i$  and in particular (2.2a). For  $T \in \mathcal{T}_H \setminus \tilde{\mathcal{R}}_H$ , the choice of  $e_i$  implies  $w_h|_{e_i} \in \mathcal{P}_{r+5}(e_i)$ ,  $\nabla w_h|_{e_i} \in \mathcal{P}_{r+4}(e_i)$ . Using this fact and (5.1), (5.2), we obtain

$$D_i(I_H w_h) = D_i w_h \text{ for all } a_i \subset \bar{T}. \tag{5.4}$$

Due to (5.4) and the unisolvence of  $\{D_i\}_{a_i \in T}$  (see the analysis in [24, 37]), we have  $(w_h - I_H w_h)|_T = 0$  and thus verify the property (2.2b).

We note that  $w_h \in W_h$  implies  $(\nabla w_h)|_{\Gamma_j}$  is constant and thus  $D_i w_h = D_i I_H w_h$  with  $|D_i| = 1$  for each boundary node  $a_i \in \Gamma_j$ . It could also be observed from (5.3), (5.2), and constancy of  $(\nabla w_h)|_{\Gamma_j}$  that  $\partial_{e_i} \partial_{e_i} I_H w_h = \partial_{e_i} \partial_{e_i} w_h$  for each boundary  $e_i \subset \Gamma_j$ . Therefore enough nodal variables vanish to enforce  $w_h = I_H w_h$  on  $\Gamma_j$  and (2.2c) is confirmed.

For the same reason above,  $\partial_{e_i^1} \partial_{e_i^2} w_h = \partial_{e_i^1} \partial_{e_i^2} I_H w_h$  for each cross derivative assigned to boundary node  $a_i \in \Gamma_j$ . Therefore enough nodal variables vanish to enforce  $\partial_n(w_h - I_H w_h)|_{\Gamma_j} = 0$ .

The interpolation estimate (2.3) directly follows from the same proof of Theorem 7.3 in [24] together with a trace inequality.  $\square$

The rest of this section is devoted to the proof of Lemma 3.1. First we present a modified Korn's inequality on each local triangle  $T \in \mathcal{T}_h$ .

**Lemma 5.1.** *Given  $\mathcal{T}_h \in \mathbb{T}$ ,  $T \in \mathcal{T}_h$  and  $v \in H^1(T, \mathbb{R}^2)$ , we have*

$$|v|_{H^1(T)} \leq C_{\text{Korn}} (\|\varepsilon(v)\|_T + h_T^{-1} \|Q_T v\|_T),$$

where  $Q_T v$  is the  $L^2$  projection of  $v$  onto  $\mathcal{RM}$ , and  $C_{\text{Korn}}$  is a constant relying only on  $\gamma_0$ .

*Proof.* The standard compactness argument (cf. [9], Thm. 11.2.16) implies

$$\|v\|_{H^1(T)} \leq C_T(\|\varepsilon(v)\|_T + \|Q_T v\|_T), \tag{5.5}$$

where  $C_T$  is a constant depending on  $T$ . It remains to estimate  $C_T$  by a homogeneity argument. Consider a reference triangle  $K$  and the affine mapping  $F_T : K \rightarrow T$  given by  $F_T(x) = B_T x + b_T$ . Define

$$\Phi(B_T) = \sup_{v \in H^1(K, \mathbb{R}^2), \|v\|_{H^1(K)}=1} \Phi_v(B_T),$$

where

$$\Phi_v(B_T) = \frac{|v \circ F_T^{-1}|_{H^1(T)}}{\|\varepsilon(v \circ F_T^{-1})\|_T + h_T^{-1} \|Q_T(v \circ F_T^{-1})\|_T}.$$

Note that  $\Phi_v$  is independent of  $b_T$ . Due to (5.5), the function  $\Phi$  is well-defined. It is straightforward to check that

$$\{\Phi_v(\cdot)\}_{v \in H^1(K, \mathbb{R}^2), \|v\|_{H^1(K)}=1}$$

is a family of equicontinuous functions on  $GL(2, \mathbb{R})$ . Hence  $\Phi$  defined by taking the supremum of this family must be continuous on  $GL(2, \mathbb{R})$ . Let  $\hat{T} = \{h_T^{-1}x : x \in T\}$  be the scaled triangle of unit size. Since  $\mathcal{T}_h$  is shape regular,  $\{B_{\hat{T}} : T \in \mathcal{T}_h\}$  is contained in a compact subset of  $GL(2, \mathbb{R})$ , see, e.g., [9]. Combining the continuity and compactness, we obtain

$$\sup_{T \in \mathcal{T}_h} \Phi(B_{\hat{T}}) = C_{\text{sup}} < \infty, \tag{5.6}$$

where  $C_{\text{sup}}$  depends on the shape regularity of  $\mathcal{T}_h$ . Therefore using a *scaling* transformation and (5.6), we obtain

$$\begin{aligned} \Phi(B_T) &= \sup_{v \in H^1(K, \mathbb{R}^2), \|v\|_{H^1(K)}=1} \frac{|v \circ F_{\hat{T}}^{-1}|_{H^1(\hat{T})}}{\|\varepsilon(v \circ F_{\hat{T}}^{-1})\|_{\hat{T}} + \|Q_{\hat{T}}(v \circ F_{\hat{T}}^{-1})\|_{\hat{T}}} \\ &\lesssim \Phi(B_{\hat{T}}) \leq C_{\text{sup}}. \end{aligned}$$

The proof is complete. □

Then we present a modified discrete Korn’s inequality on a triangle  $T$ .

**Lemma 5.2.** *Let  $\mathcal{T}_h, \mathcal{T}_H \in \mathbb{T}$  with  $\mathcal{T}_H \leq \mathcal{T}_h$ . For any  $T \in \mathcal{T}_H$ , let  $\mathcal{T}_h(T) = \{T' \in \mathcal{T}_h : T' \subset T\}$  and  $\mathcal{E}_h(\overset{\circ}{T}) = \{e \in \mathcal{E}_h : e \subset T, e \not\subseteq \partial T\}$ . Then for  $v_h \in U_h|_T$ , we have*

$$h_T^{-2} \|v_h - Q_T v_h\|_T^2 \lesssim \sum_{T' \in \mathcal{T}_h(T)} \|\varepsilon(v_h)\|_{T'}^2 + \sum_{e \in \mathcal{E}_h(\overset{\circ}{T})} h_e^{-1} \|[[v_h]]\|_e^2.$$

*Proof.* Let  $w = v_h - Q_T v_h$  and  $|w|_{H_h^1(T)}^2 := \sum_{T' \in \mathcal{T}_h(T)} |w|_{H^1(T')}^2$ . Let

$$S_h(T) = \{\tilde{v} \in C^0(T) : \tilde{v}|_{T'} \in \mathcal{P}_{r+2}(T') \text{ for } T' \in \mathcal{T}_h(T)\}$$

be the usual Lagrange element space of degree  $r+2$ . Following the analysis in [8,9,31], we construct a continuous piecewise polynomial function  $Ew \in S_h(T)$  by setting the nodal value as

$$Ew(x) = \frac{1}{\#\omega_{h,x}} \sum_{T' \in \omega_{h,x}} (w|_{T'})(x),$$

where  $x$  is a Lagrange node for the space  $S_h(T)$  and  $\omega_{h,x} = \{T' \in \mathcal{T}_h(T) : x \in T'\}$ . An elementary estimate shows that

$$h_T^{-2} \|w - Ew\|_T^2 + |w - Ew|_{H_h^1(T)}^2 \lesssim \sum_{e \in \mathcal{E}_h(\hat{T})} h_e^{-1} \|\llbracket w \rrbracket_e\|^2, \tag{5.7}$$

see, e.g., the proof of Lemma 10.6.6 in [9] and Lemma 2.8 in [31]. For the continuous function  $Ew$ , the Poincaré inequality implies

$$\|Ew - Q_T Ew\|_T \lesssim h_T |Ew|_{H^1(T)}, \tag{5.8}$$

Using (5.7), (5.8), the triangle inequality, and  $Q_T w = 0$ , we have

$$\begin{aligned} \|w\|_T^2 &\lesssim \|w - Ew\|_T^2 + \|Q_T(w - Ew)\|_T^2 + \|Ew - Q_T Ew\|_T^2 \\ &\lesssim h_T^2 \sum_{e \in \mathcal{E}_h(\hat{T})} h_e^{-1} \|\llbracket w \rrbracket_e\|^2 + h_T^2 |Ew|_{H^1(T)}^2. \end{aligned} \tag{5.9}$$

It remains to estimate  $|Ew|_{H^1(T)}$ . Lemma 5.1 implies

$$|Ew|_{H^1(T)} \lesssim \|\varepsilon(Ew)\|_T + h_T^{-1} \|Q_T Ew\|_T. \tag{5.10}$$

It then follows from the triangle inequality, (5.10),  $\varepsilon(Q_T v_h) = 0$ , and  $Q_T(w) = 0$  that

$$\begin{aligned} |Ew|_{H^1(T)}^2 &\lesssim \sum_{T' \in \mathcal{T}_h(T)} (\|\varepsilon(w)\|_{T'}^2 + \|\varepsilon(w - Ew)\|_{T'}^2) + h_T^{-2} \|Q_T(Ew - w)\|_T^2 \\ &\leq \sum_{T' \in \mathcal{T}_h(T)} \|\varepsilon(v_h)\|_{T'}^2 + |w - Ew|_{H_h^1(T)}^2 + h_T^{-2} \|w - Ew\|_T^2. \end{aligned} \tag{5.11}$$

Combining (5.9), (5.11), (5.7), and  $\llbracket w \rrbracket_e = \llbracket v_h \rrbracket_e$  completes the proof. □

For  $v_h \in U_h$ , define the mesh-dependent norm

$$|v_h|_{1,h} := \left( \sum_{T \in \mathcal{T}_h} \|\varepsilon(v_h)\|_T^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|\llbracket v_h \rrbracket_e\|^2 \right)^{\frac{1}{2}}.$$

It has been shown in [19] that the following discrete inf-sup condition holds:

$$|v_h|_{1,h} \lesssim \sup_{\tau_h \in \Sigma_h^{\text{HZ}}} \frac{(\text{div } \tau_h, v_h)}{\|\tau_h\|} \text{ for all } v_h \in U_h. \tag{5.12}$$

With the above preparation, we are able to prove Lemma 3.1.

*Proof of Lemma 3.1.* Using the inf-sup condition (5.12) and the inclusion  $\Sigma_h^{\text{HZ}} \subset \Sigma_h$ , we obtain

$$|v_h|_{1,h} \lesssim \sup_{\tau_h \in \Sigma_h} \frac{(\text{div } \tau_h, v_h)}{\|\tau_h\|} = \sup_{\tau_h \in \Sigma_h} \frac{(\mathbb{A} \tau_h, \varepsilon_{\mathbb{C}}^h(v_h))}{\|\tau_h\|}. \tag{5.13}$$

It then follows from (5.13) and  $\|\tau_h\|_{\mathbb{A}} \lesssim \|\tau_h\|$  that

$$|v_h|_{1,h} \lesssim \|\varepsilon_{\mathbb{C}}^h(v_h)\|_{\mathbb{A}}. \tag{5.14}$$

Combining it with Lemma 5.2, we have

$$\begin{aligned} \sum_{T \in \mathcal{T}_H} h_T^{-2} \|v_h - Q_H v_h\|_T^2 &\lesssim \sum_{T \in \mathcal{T}_H} \left( \sum_{T' \in \mathcal{T}_h(T)} \|\varepsilon(v_h)\|_{T'}^2 + \sum_{e \in \mathcal{E}_h(\hat{T})} h_e^{-1} \|\llbracket v_h \rrbracket_e\|^2 \right) \\ &\lesssim |v_h|_{1,h}^2 \lesssim \|\varepsilon_{\mathbb{C}}^h(v_h)\|_{\mathbb{A}}^2, \end{aligned}$$

which completes the proof. □

6. IMPLEMENTATION AND NUMERICAL EXPERIMENT

The method (1.4) can be implemented using the hybridization technique. Consider the multiplier space

$$M_h = \{ \mu_h : \mu_h|_e \in \mathcal{P}_{r+3}(e, \mathbb{R}^2) \text{ for all } e \in \mathcal{E}_h^o \},$$

and the broken discrete stress space

$$\Sigma_h^{-1} = \{ \tau_h \in L^2(\Omega, \mathbb{S}) : \tau_h|_T \in \mathcal{P}_{r+3}(T) \text{ for all } T \in \mathcal{T}_h \}.$$

The hybridized mixed method seeks  $(\tilde{\sigma}_h, \tilde{u}_h, \lambda_h) \in \Sigma_h^{-1} \times U_h \times M_h$  such that

$$\begin{aligned} (\mathbb{A}\tilde{\sigma}_h, \tau_h) + \sum_{T \in \mathcal{T}_h} (\operatorname{div} \tau_h, \tilde{u}_h)_T + \sum_{e \in \mathcal{E}_h^o} \int_e \lambda_h \cdot \llbracket \tau_h \rrbracket n_e \, ds &= \langle \tau_h n, g_D \rangle_{\Gamma_D}, \\ \sum_{T \in \mathcal{T}_h} (\operatorname{div} \tilde{\sigma}_h, v_h)_T &= (f, v_h), \\ \sum_{e \in \mathcal{E}_h^o} \int_e \mu_h \cdot \llbracket \tilde{\sigma}_h \rrbracket n_e \, ds &= 0, \end{aligned} \tag{6.1}$$

for all  $\tau \in \Sigma_h^{-1}$ ,  $v_h \in U_h$ ,  $\mu_h \in M_h$ . In fact, (6.1) is a hybridized version of (1.4), *i.e.*,  $\tilde{\sigma}_h = \sigma_h, \tilde{u}_h = u_h$ , see [1, 25]. Because  $\Sigma_h^{-1}$ ,  $U_h$ ,  $M_h$  are completely broken, it is straightforward to construct their local basis. In matrix notation, (6.1) reads

$$\begin{pmatrix} A & B \\ B^\top & O \end{pmatrix} \begin{pmatrix} X \\ \Lambda \end{pmatrix} = \begin{pmatrix} F \\ O \end{pmatrix}, \tag{6.2}$$

where  $O$  is a zero matrix or vector,  $X$  and  $\Lambda$  are vectors corresponding to the coordinates of  $(\tilde{\sigma}_h, \tilde{u}_h)$  and  $\lambda_h$ , respectively.

Due to the discontinuity of  $\Sigma_h^{-1}$  and  $U_h$ , the matrix  $A$  is block diagonal and easily invertible. Hence solving (6.2) is equivalent to solving the smaller Schur complement system

$$B^\top A^{-1} B \Lambda = B^\top A^{-1} F. \tag{6.3}$$

Here  $B^\top A^{-1} B$  is a sparse and positive semi-definite matrix and the size of  $B^\top A^{-1} B$  is much smaller than (6.1) or (1.4). However, the Schur complement  $B^\top A^{-1} B$  has a small kernel provided  $\mathcal{T}_h$  has singular vertices and/or pure traction boundary condition ( $\Gamma_D = \emptyset$ ) is considered. The key point is that such kernel could be easily resolved by iterative methods such as the preconditioned conjugate gradient method. An optimal preconditioner for the Schur complement system (6.3) is presented in [25].

In the experiment, let  $\Omega = [-1, 1]^2 \setminus ([0, 1] \times [-1, 0])$  be the L-shaped domain. Let  $(r, \theta)$  be the polar coordinate with respect to the origin, where  $0 \leq \theta \leq \omega = \frac{3\pi}{2}$ . Let

$$\begin{aligned} \Phi_1(\theta) &= \begin{pmatrix} ((z+2)(\lambda+\mu) + 4\mu) \sin(z\theta) - z(\lambda+\mu) \sin((z-2)\theta) \\ z(\lambda+\mu)(\cos(z\theta) - \cos((z-2)\theta)) \end{pmatrix}, \\ \Phi_2(\theta) &= \begin{pmatrix} z(\lambda+\mu)(\cos((z-2)\theta) - \cos(z\theta)) \\ -((2-z)(\lambda+\mu) + 4\mu) \sin(z\theta) - z(\lambda+\mu) \sin((z-2)\theta) \end{pmatrix}, \end{aligned}$$

and

$$\begin{aligned} \Phi(\theta) &= \{ z(\lambda+\mu) \sin((z-2)\omega) + ((2-z)(\lambda+\mu) + 4\mu) \sin(z\omega) \} \Phi_1(\theta) \\ &\quad - z(\lambda+\mu)(\cos((z-2)\omega) - \cos(z\omega)) \Phi_2(\theta). \end{aligned}$$

where  $z \in (0, 1)$  is a root of  $(\lambda + 3\mu)^2 \sin^2(z\omega) = (\lambda + \mu)^2 z^2 \sin^2(\omega)$ . The most singular part of the solution to (1.2) behaves like  $r^z \Phi(\theta)$  in the neighborhood of  $(0, 0)$ , see, *e.g.*, [26]. Therefore we choose

$$u(r, \theta) = \frac{1}{(\lambda + \mu)^2} (x_1^2 - 1)(x_2^2 - 1) r^z \Phi(\theta)$$



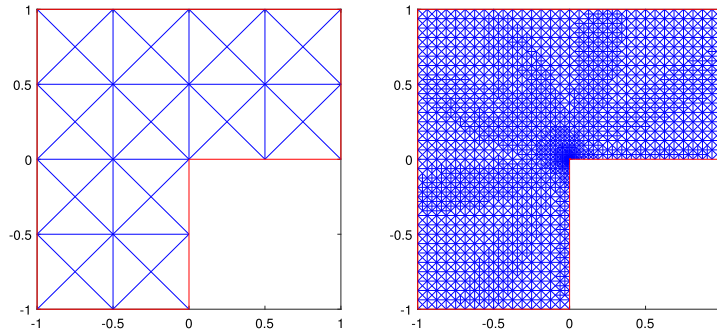


FIGURE 1. *Left*: initial grid. *Right*: adaptive grid, 5290 elements.

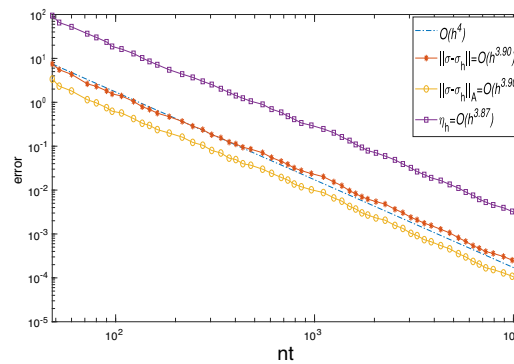


FIGURE 2. Error curve.

as the exact solution in the test problem. The boundary condition is based on pure displacement ( $\Gamma_N = \emptyset$ ). The Lamé constants are  $\lambda = 10^4$  and  $\mu = 1$ . The method (1.4) or (6.1) is implemented using the package iFEM [17] in Matlab 2019a. We start with the initial mesh in Figure 1 and set the marking parameter  $\theta = 0.3$ . The algebraic system (6.3) is solved by the conjugate gradient method preconditioned by the incomplete Cholesky decomposition. Numerical results are presented in Figure 2, where  $nt$  denotes the number of triangles.

It can be observed from Figure 1(right) that the adaptive Algorithm 4.1 captures the corner singularity. Figure 2 shows that Algorithm 4.1 has optimal and robust rate of convergence with respect to very large Lamé constant  $\lambda$  starting from coarse initial grid, which validates our convergence and complexity analysis.

*Acknowledgements.* The author would like to thank Dr. Shihua Gong for generously sharing his Matlab code and comments on iterative methods.

## REFERENCES

- [1] D.N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO: M2AN* **19** (1985) 7–32.
- [2] D.N. Arnold and R. Winther, Mixed finite elements for elasticity. *Numer. Math.* **92** (2002) 401–419.
- [3] D.N. Arnold, J. Douglas, Jr. and C.P. Gupta, A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.* **45** (1984) 1–22.
- [4] D.N. Arnold, G. Awanou and R. Winther, Finite elements for symmetric tensors in three dimensions. *Math. Comput.* **77** (2008) 1229–1251.
- [5] E. Bänsch, Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Eng.* **3** (1991) 181–191.

- [6] R. Becker and S. Mao, An optimally convergent adaptive mixed finite element method. *Numer. Math.* **111** (2008) 35–54.
- [7] P. Binev, W. Dahmen and R. DeVore, Adaptive finite element methods with convergence rates. *Numer. Math.* **97** (2004) 219–268.
- [8] S.C. Brenner, Korn’s inequalities for piecewise  $H^1$  vector fields. *Math. Comput.* **73** (2003) 1067–1087.
- [9] S.C. Brenner and L. Ridgway Scott, The mathematical theory of finite element methods, 3 edition. In: Vol. 35 of Texts in Applied Mathematics, 15. Springer, New York (2008).
- [10] C. Carstensen, A unifying theory of a posteriori finite element error control. *Numer. Math.* **100** (2005) 617–637.
- [11] C. Carstensen and G. Dolzmann, A posteriori error estimates for mixed FEM in elasticity. *Numer. Math.* **81** (1998) 187–209.
- [12] C. Carstensen and R.H.W. Hoppe, Error reduction and convergence for an adaptive mixed finite element method. *Math. Comput.* **75** (2006) 1033–1042.
- [13] C. Carstensen and H. Rabus, The adaptive nonconforming FEM for the pure displacement problem in linear elasticity is optimal and robust. *SIAM J. Numer. Anal.* **50** (2012) 1264–1283.
- [14] C. Carstensen, M. Feischl, M. Page and D. Praetorius, Axioms of adaptivity. *Comput. Math. Appl.* **67** (2014) 1195–1253.
- [15] C. Carstensen, D. Gallistl and J. Gedicke, Residual-based a posteriori error analysis for symmetric mixed Arnold-Winther FEM. *Numer. Math.* **142** (2019) 205–234.
- [16] J.M. Cascon, C. Kreuzer, R.H. Nochetto and K.G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46** (2008) 2524–2550.
- [17] L. Chen, iFEM: an innovative finite element method package in Matlab. University of California Irvine, Technical report (2009).
- [18] L. Chen, M. Holst and J. Xu, Convergence and optimality of adaptive mixed finite element methods. *Math. Comput.* **78** (2009) 35–53.
- [19] L. Chen, J. Hu and X. Huang, Fast auxiliary space preconditioners for linear elasticity in mixed form. *Math. Comput.* **78** (2018) 1601–1633.
- [20] L. Chen, J. Hu, X. Huang and H. Man, Residual-based a posteriori error estimates for symmetric conforming mixed finite elements for linear elasticity problems. *Sci. China Math.* **61** (2018) 973–992.
- [21] L. Diening, C. Kreuzer and R. Stevenson, Instance optimality of the adaptive maximum strategy. *Found. Comput. Math.* **16** (2016) 33–68.
- [22] W. Dörfler, A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.
- [23] M. Feischl, T. Führer and D. Praetorius, Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. *SIAM J. Numer. Anal.* **52** (2014) 601–625.
- [24] V. Girault and L.R. Scott, Hermite interpolation of nonsmooth functions preserving boundary conditions. *Math. Comput.* **71** (2002) 1043–1074.
- [25] S. Gong, S. Wu and J. Xu, New hybridized mixed methods for linear elasticity and optimal multilevel solvers. *Numer. Math.* **141** (2019) 569–604.
- [26] P. Grisvard, Singularities in boundary value problems. In: Vol. 22 of Research in Applied Mathematics. Springer-Verlag, Berlin (1992).
- [27] M. Holst, Y. Li, A. Mihalik and R. Szypowski, Convergence and optimality of adaptive mixed methods for Poisson’s equation in the FEEC framework. *J. Comput. Math.* **38** (2020) 748–767.
- [28] J. Hu, Finite element approximations of symmetric tensors on simplicial grids in  $\mathbb{R}^n$ : the higher order case. *J. Comput. Math.* **33** (2015) 283–296.
- [29] J. Hu and R. Ma, Partial relaxation of  $C^0$  vertex continuity of stresses of conforming mixed finite elements for the elasticity problem. *Comput. Methods Appl. Math.* **21** (2021) 89–108.
- [30] J. Hu and S. Zhang, A family of conforming mixed finite elements for linear elasticity on triangular grids. Preprint [arXiv:1406.7457](https://arxiv.org/abs/1406.7457) (2014).
- [31] J. Huang and Y. Xu, Convergence and complexity of arbitrary order adaptive mixed element methods for the poisson equation. *Sci. China Math.* **55** (2012) 1083–1098.
- [32] Y. Li, Some convergence and optimality results of adaptive mixed methods in finite element exterior calculus. *SIAM J. Numer. Anal.* **57** (2019) 2019–2042.
- [33] Y. Li, Quasi-optimal adaptive mixed finite element methods for controlling natural norm errors. *Math. Comput.* **90** (2021) 565–593.
- [34] Y. Li and L. Zikatanov, Nodal auxiliary a posteriori error estimates. Preprint [arXiv:2010.06774](https://arxiv.org/abs/2010.06774) (2020).
- [35] M. Lonsing and R. Verfürth, A posteriori error estimators for mixed finite element methods in linear elasticity. *Numer. Math.* **97** (2004) 757–778.
- [36] W.F. Mitchell, A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Softw.* **15** (1989) 326–347.
- [37] J. Morgan and R. Scott, A nodal basis for  $C^1$  piecewise polynomials of degree  $n \geq 5$ . *Math. Comput.* **29** (1975) 736–740.
- [38] P. Morin, R.H. Nochetto and K.G. Siebert, Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.* **38** (2000) 466–488.
- [39] R. Stevenson, Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7** (2007) 245–269.
- [40] R. Stevenson, The completion of locally refined simplicial partitions created by bisection. *Math. Comput.* **77** (2008) 227–241.
- [41] R. Verfürth, A posteriori error estimation techniques for finite element methods. In: Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford (2013).

## Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

**Please help to maintain this journal in open access!**

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org)

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/maths-s2o-programme>