MATHEMATICS OF COMPUTATION
Volume 87, Number 309, January 2018, Pages 69–93
http://dx.doi.org/10.1090/mcom/3249
Article electronically published on May 1, 2017

AN HDG METHOD FOR LINEAR ELASTICITY WITH STRONG SYMMETRIC STRESSES

WEIFENG QIU, JIGUANG SHEN, AND KE SHI

ABSTRACT. This paper presents a new hybridizable discontinuous Galerkin (HDG) method for linear elasticity on general polyhedral meshes, based on a strong symmetric stress formulation. The key feature of this new HDG method is the use of a special form of the numerical trace of the stresses, which makes the error analysis different from the projection-based error analyzes used for most other HDG methods. For arbitrary polyhedral elements, we approximate the stress by using polynomials of degree $k \geq 1$ and the displacement by using polynomials of degree $k+1$. In contrast, to approximate the numerical trace of the displacement on the faces, we use polynomials of degree k only. This allows for a very efficient implementation of the method, since the numerical trace of the displacement is the only globally-coupled unknown, but does not degrade the convergence properties of the method. Indeed, we prove optimal orders of convergence for both the stresses and displacements on the elements. In the almost incompressible case, we show the error of the stress is also optimal in the standard L^2-norm. These optimal results are possible thanks to a special superconvergence property of the numerical traces of the displacement, and thanks to the use of the crucial elementwise Korn's inequality. Several numerical results are presented to support our theoretical findings in the end.

1. INTRODUCTION

In this paper, we introduce a new hybridizable discontinuous Galerkin (HDG) method for the system of linear elasticity:

$$\mathcal{A}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) = 0 \quad \text{in } \Omega \subset \mathbb{R}^3, \tag{1.1a}$$

$$\nabla \cdot \underline{\boldsymbol{\sigma}} = \boldsymbol{f} \quad \text{in } \Omega, \tag{1.1b}$$

$$\boldsymbol{u} = \boldsymbol{g} \quad \text{on } \partial\Omega. \tag{1.1c}$$

Here, the displacement is denoted by the vector field $\boldsymbol{u} : \Omega \to \mathbb{R}^3$. The strain tensor is represented by $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) := \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top)$. The stress tensor is represented by $\underline{\boldsymbol{\sigma}} : \Omega \to \boldsymbol{S}$, where $\boldsymbol{S}$ denotes the set of all symmetric matrices in $\mathbb{R}^{3\times 3}$. The compliance tensor $\mathcal{A}$ is assumed to be a bounded, symmetric, positive definite tensor over $\boldsymbol{S}$. The body force $\boldsymbol{f}$ lies in $\boldsymbol{L}^2(\Omega)$, the displacement of the boundary $\boldsymbol{g}$ is a function in $\boldsymbol{H}^{1/2}(\partial\Omega)$ and Ω is a polyhedral domain.

Received by the editor August 14, 2014 and, in revised form, November 15, 2015, February 16, 2016, and August 23, 2016.

2010 *Mathematics Subject Classification.* Primary 65N30, 65L12.

Key words and phrases. Hybridizable, discontinuous Galerkin, superconvergence, linear elasticity.

In general, there are two approaches to design mixed finite element methods for linear elasticity. The first approach is to enforce the symmetry of the stress tensor weakly ([4, 5, 11, 17, 25, 28, 32, 33, 36]). In this category, is included the HDG method considered in [22]. The other approach is to exactly enforce the symmetry of the approximate stresses. The methods considered in [1–3, 7, 8, 21, 26, 30, 35, 37, 38] belong to the second category, and so does the contribution of this paper. In general, the methods in the first category are easier to implement. On the other hand, the methods in the second category preserve the balance of angular momentum strongly and have less degrees of freedom. Next, we compare our HDG method with several methods of the second category.

In [21], an LDG method using strongly symmetric stresses (for isotropic linear elasticity) was introduced and proved to yield convergence properties that remain unchanged when the material becomes incompressible; simplexes and polynomial approximations or degree k in all variables were used. However, as all LDG methods for second-order elliptic problems, although the displacement converges with order $k+1$, the strain and pressure converge suboptimally with order k. Also, the method cannot be hybridized. Stress finite elements satisfying both strong symmetry and $H(\text{div})$-conformity are introduced in [1, 2]. The main drawback of these methods is that they have too many degrees of freedom of stress elements and hybridization is not available for them (see detailed description in [28]). In [3, 7, 8, 26, 30, 35, 37, 38], non-conforming methods using symmetric stress elements are introduced. But, methods in [3, 7, 8, 30, 37, 38] use low order finite element spaces only (most of them are restricted to rectangular or cubical meshes except [3, 7]). In [26], a family of simplicial elements (one for each $k \geq 1$) are developed in both two and three dimensions. (The degrees of freedom of $P_{k+1}(\boldsymbol{S}, K)$ were studied in [26] and then used to design the projection operator $\Pi^{(\text{div},\boldsymbol{S})}$ in [27]). However, the convergence rate of stress is suboptimal. The first HDG method for linear and nonlinear elasticity was introduced in [34, 35]; see also the related HDG method proposed in [39]. These methods also use simplexes and polynomial approximations of degree k in all variables. For general polyhedral elements, this method was recently analyzed in [23] where it was shown that the method converges optimally in the displacement with order $k+1$, but with the suboptimal order of $k+1/2$ for the pressure and the stress. For $k=1$, these orders of convergence were numerically shown to be sharp for triangular elements. In this paper, we prove that by enriching the local stress space to be polynomials of degree no more than $k+1$, and by using a modified numerical trace, we are able to obtain optimal order of convergence for all unknowns. In addition, this analysis is valid for general polyhedral meshes. To the best of our knowledge, this is so far the only result which has optimal accuracy with general polyhedral triangulations for linear elasticity problems.

Like many hybrid methods, our HDG method provides approximation to stress and displacement in each element and trace of displacement along interfaces of meshes. In general, the corresponding finite element spaces are $\underline{\boldsymbol{V}}_h, \boldsymbol{W}_h, \boldsymbol{M}_h$, which are defined to be

$$\begin{aligned}
\underline{\boldsymbol{V}}_h &= \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\Omega) : \quad \underline{\boldsymbol{v}}|_K \in \underline{\boldsymbol{V}}(K) \quad \forall\, K \in \mathcal{T}_h\},\\
\boldsymbol{W}_h &= \{\boldsymbol{\omega} \in \boldsymbol{L}^2(\Omega) : \quad \boldsymbol{\omega}|_K \in \boldsymbol{W}(K) \quad \forall\, K \in \mathcal{T}_h\},\\
\boldsymbol{M}_h &= \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \quad \boldsymbol{\mu}|_F \in \boldsymbol{M}(F) \quad \forall\, F \in \mathcal{E}_h\}.
\end{aligned}$$

Here $\mathcal{T}_h$ denotes a triangulation of the domain Ω and $\mathcal{E}_h$ is the set of all faces F of all elements $K \in \mathcal{T}_h$. The spaces $\underline{\boldsymbol{V}}(K), \boldsymbol{W}(K), \boldsymbol{M}(F)$ are called the *local spaces* which are defined on each element/face. In Table 1 we list several choices of local spaces for different methods. In this paper, our choice of the local spaces is defined as:

$$\underline{\boldsymbol{V}}(K) = \underline{\boldsymbol{P}}_k(\boldsymbol{S}, K), \quad \boldsymbol{W}(K) = \boldsymbol{P}_{k+1}(K), \quad \boldsymbol{M}(F) = \boldsymbol{P}_k(F).$$

Here, the space of vector-valued functions defined on D whose entries are polynomials of total degree k is denoted by $\boldsymbol{P}_k(D)$ $(k \geq 1)$. Similarly, $\underline{\boldsymbol{P}}_k(\boldsymbol{S}, K)$ denotes the space of symmetric-valued functions defined on K whose entries are polynomials of total degree k. In addition, our method allows $\mathcal{T}_h$ to be any conforming polyhedral triangulation of Ω.

Note the fact that the only globally-coupled degrees of freedom are those of the numerical trace of displacement along $\mathcal{E}_h$, renders the method efficiently implementable. However, the fact that the polynomial degree of the approximate numerical traces of the displacement is one *less* than that of the approximate displacement inside the elements, might cause a degradation in the approximation properties of the displacement. However, this unpleasant situation is avoided altogether by taking a special form of the numerical trace of the stresses inspired on the choice taken in [29] in the framework of diffusion problems. This choice allows for a special superconvergence of part of the numerical traces of the stresses which, in turn, guarantees that, for $k \geq 1$, the L^2-order of convergence for the stress is $k+1$ and that of the displacement $k+2$. So, we obtain optimal convergence for both stress and displacement for general polyhedral elements. Let us mention that the approach of error analysis of our HDG method is different from the traditional projection-based error analysis in [19, 20, 22] in three aspects. First, here, we use simple L^2-projections, not the numerical trace-tailored projections typically used for the analysis of other HDG methods. Second, we take the stabilization parameter to be of order $1/h$ instead of of order one. And finally, we use an elementwise Korn's inequality (Lemma 4.1) to deal with the symmetry of the stresses.

We notice that mixed methods in [17, 25] and HDG methods in [22] also achieve optimal convergence for stress and superconvergence for displacement by post processing. However, there are two disadvantages regarding implementation. First, these methods enforce the stress symmetry weakly, which means that they have a much larger space for the stress. In additon, these methods usually need to add matrix bubble functions ($\underline{\boldsymbol{\delta V}}$ in [17]) into their stress elements in order to obtain optimal approximations. In fact, the construction of such bubbles on general polyhedral elements is still an open problem. In contrast, our method avoids using matrix bubble functions but only uses simple polynomial spaces of degree $k, k+1$. In Table 1, we compare methods which use $\boldsymbol{M}_h$ for approximating the trace of displacement $\widehat{\boldsymbol{u}}_h$ on $\mathcal{E}_h$. There, $\boldsymbol{u}_h^\star$ is a post-processed numerical solution of displacement.

The remainder of this paper is organized as follows. In Section 2, we introduce our HDG method and present our a priori error estimates. In Section 3, we give a characterization of the HDG method and show the global matrix is symmetric and positive definite. In Section 4, we give the elementwise Korn's inequality in Lemma 4.1, then provide a detailed proof of the a priori error estimates. In Section 5, we present several numerical examples in order to illustrate and test our method.

TABLE 1. Orders of convergence for methods for which $\widehat{\boldsymbol{u}}_h \in \boldsymbol{M}(F) = \boldsymbol{P}_k(F),\ k \geq 1$, and K is a tetrahedron.

method	$\underline{\boldsymbol{V}}(K)$	$\boldsymbol{W}(K)$	$\Vert\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\Vert_{\mathcal{T}_h}$	$\Vert \boldsymbol{u} - \boldsymbol{u}_h\Vert_{\mathcal{T}_h}$	$\Vert\boldsymbol{u} - \boldsymbol{u}_h^\star\Vert_{\mathcal{T}_h}$
AFW[5]	$\underline{\boldsymbol{P}}_k(\mathbb{R}^{3\times3}, K)$	$\boldsymbol{P}_{k-1}(K)$	k	k	-
CGG[17]	$\underline{\mathbf{RT}}_k(K) + \underline{\boldsymbol{\delta V}}$	$\boldsymbol{P}_k(K)$	$k+1$	$k+1$	$k+2$
GG[25]	$\underline{\boldsymbol{P}}_k(\mathbb{R}^{3\times3}, K) + \underline{\boldsymbol{\delta V}}$	$\boldsymbol{P}_{k-1}(K)$	$k+1$	k	$k+1$
CS[22]	$\underline{\boldsymbol{P}}_k(\mathbb{R}^{3\times3}, K) + \underline{\boldsymbol{\delta V}}$	$\boldsymbol{P}_k(K)$	$k+1$	$k+1$	$k+2$
GG[26]	$\underline{\boldsymbol{P}}_{k+1}(\boldsymbol{S}, K)$	$\boldsymbol{P}_k(K)$	k	$k+1$	-
HDG-S	$\underline{\boldsymbol{P}}_k(\boldsymbol{S}, K)$	$\boldsymbol{P}_{k+1}(K)$	$k+1$	$k+2$	-

2. Main results

In this section we first present the method in details and then show the main results for the error estimates.

2.1. **The HDG formulation with strong symmetry.**

Let us begin by introducing some notation and conventions. We adapt to our setting the notation used in [20]. Let $\mathcal{T}_h$ denote a conforming triangulation of Ω made of shape-regular polyhedral elements K. We recall that $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$, and $\mathcal{E}_h$ denotes the set of all faces F of all elements. We denote by $\mathcal{F}(K)$ the set of all faces F of the element K. We also use the standard notation to denote scalar, vector and tensor spaces. Thus, if $D(K)$ denotes a space of scalar-valued functions defined on K, the corresponding space of vector-valued functions is $\boldsymbol{D}(K) := [D(K)]^d$ and the corresponding space of matrix-valued functions is $\underline{\boldsymbol{D}}(K) := [D(K)]^{d\times d}$. Finally, $\underline{\boldsymbol{D}}(\boldsymbol{S}, K)$ denotes the symmetric subspace of $\underline{\boldsymbol{D}}(K)$.

The methods we consider seek an approximation $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ to the exact solution $(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h})$ in the finite dimensional space $\underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h \subset \underline{\boldsymbol{L}}^2(\boldsymbol{S}, \Omega) \times \boldsymbol{L}^2(\Omega) \times \boldsymbol{L}^2(\mathcal{E}_h)$ given by

$$\underline{\boldsymbol{V}}_h = \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\boldsymbol{S}, \Omega) : \quad \underline{\boldsymbol{v}}|_K \in \underline{\boldsymbol{P}}_k(\boldsymbol{S}, K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.1a}$$

$$\boldsymbol{W}_h = \{\boldsymbol{\omega} \in \boldsymbol{L}^2(\Omega) : \quad \boldsymbol{\omega}|_K \in \boldsymbol{P}_{k+1}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.1b}$$

$$\boldsymbol{M}_h = \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \quad \boldsymbol{\mu}|_F \in \boldsymbol{P}_k(F) \quad \forall\, F \in \mathcal{E}_h\}. \tag{2.1c}$$

Here $P_k(D)$ denotes the standard space of polynomials of degree no more than k on D. Here we require $k \geq 1$.

The numerical approximation $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ can now be defined as the solution of the following system:

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}_h\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,,\, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle\widehat{\boldsymbol{u}}_h\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{2.2a}$$

$$(\underline{\boldsymbol{\sigma}}_h\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = -(\boldsymbol{f}\,,\, \boldsymbol{\omega})_{\mathcal{T}_h}, \tag{2.2b}$$

$$\langle\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{2.2c}$$

$$\langle\widehat{\boldsymbol{u}}_h\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega} = \langle\boldsymbol{g}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega}, \tag{2.2d}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h$, where

$$\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n} = \underline{\boldsymbol{\sigma}}_h\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h) \quad \text{on } \partial\mathcal{T}_h. \tag{2.2e}$$

In fact, in Christoph Lehrenfeld's thesis, the author defines the numerical flux in this way for diffusion problems (see Remark 1.2.4 in [29]). This method was then analyzed for diffusion recently in [31]. Here, $\boldsymbol{P_M}$ denotes the standard L^2-orthogonal projection from $\boldsymbol{L}^2(\mathcal{E}_h)$ onto $\boldsymbol{M}_h$. We write $(\underline{\boldsymbol{\eta}}\,,\,\underline{\boldsymbol{\zeta}})_{\mathcal{T}_h} := \sum_{i,j=1}^n (\underline{\boldsymbol{\eta}}_{i,j}\,,\,\underline{\boldsymbol{\zeta}}_{i,j})_{\mathcal{T}_h}$, $(\boldsymbol{\eta}\,,\,\boldsymbol{\zeta})_{\mathcal{T}_h} := \sum_{i=1}^n (\eta_i\,,\,\zeta_i)_{\mathcal{T}_h}$, and $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(\eta,\zeta)_K$, where $(\eta,\zeta)_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^n$. Similarly, we write $\langle\boldsymbol{\eta}\,,\,\boldsymbol{\zeta}\rangle_{\partial\mathcal{T}_h} := \sum_{i=1}^n \langle\eta_i\,,\,\zeta_i\rangle_{\partial\mathcal{T}_h}$ and $\langle\eta\,,\,\zeta\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle\eta\,,\,\zeta\rangle_{\partial K}$, where $\langle\eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^{n-1}$.

The parameter τ in (2.2e) is called the *stabilization parameter.* In this paper, we assume it is a fixed positive number on all faces. It is worth mentioning that the numerical trace (2.2e) is defined slightly different from the usual HDG setting; see [20]. Namely, in the definition, we use $\boldsymbol{P_M u}_h$ instead of $\boldsymbol{u}_h$. Indeed, this is a crucial modification in order to get the error estimate. An intuitive explanation is that we want to preserve the strong continuity of the flux across the interfaces. Without the projection $\boldsymbol{P_M}$, by (2.2c) the normal component of $\widehat{\underline{\boldsymbol{\sigma}}}_h$ is only weakly continuous across the interfaces.

2.2. **A priori error estimates.** To state our main result, we need to introduce some notation. We define

$$\|\underline{\boldsymbol{v}}\|_{L^2(\mathcal{A},\Omega)} = \sqrt{(\mathcal{A}\underline{\boldsymbol{v}}, \underline{\boldsymbol{v}})_\Omega}, \quad \forall \underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\boldsymbol{S}, \Omega).$$

We use $\|\cdot\|_{s,D}, |\cdot|_{s,D}$ to denote the usual norm and seminorm on the Sobolev space $H^s(D)$. We discard the first index s if $s = 0$. A differential operator with a subindex h means it is defined on each element $K \in \mathcal{T}_h$. Similarly, the norm $\|\cdot\|_{s,\mathcal{T}_h}$ is the discrete norm defined as $\|\cdot\|_{s,\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\|\cdot\|_{s,K}$. Finally, we need an elliptic regularity assumption stated as follows. Let $(\boldsymbol{\phi}, \underline{\boldsymbol{\psi}}) \in \boldsymbol{H}^2(\Omega) \times \underline{\boldsymbol{H}}^1(\Omega)$ be the solution of the adjoint problem:

$$\mathcal{A}\underline{\boldsymbol{\psi}} - \underline{\epsilon}(\boldsymbol{\phi}) = 0 \quad \text{in } \Omega, \tag{2.3a}$$

$$\nabla\cdot\underline{\boldsymbol{\psi}} = \boldsymbol{e_u} \quad \text{in } \Omega, \tag{2.3b}$$

$$\boldsymbol{\phi} = 0 \quad \text{on } \partial\Omega. \tag{2.3c}$$

We assume the solution $(\boldsymbol{\phi}, \underline{\boldsymbol{\psi}})$ has the following elliptic regularity property:

$$\|\underline{\boldsymbol{\psi}}\|_{1,\Omega} + \|\boldsymbol{\phi}\|_{2,\Omega} \le C_{reg}\|\boldsymbol{e_u}\|_\Omega, \tag{2.4}$$

The assumption holds in the case of planar elasticity with scalar coefficients on a convex domain; see [9].

We are now ready to state our main result.

Theorem 2.1. *If the meshes are quasi-uniform and* $\tau = \mathcal{O}(\frac{1}{h})$*, then we have*

$$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{L^2(\mathcal{A},\Omega)} \le Ch^s(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}), \tag{2.5}$$

for all $1 \le s \le k+1$*. Moreover, if the elliptic regularity property* (2.4) *holds, then we have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_\Omega \le Ch^{s+1}(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}), \tag{2.6}$$

for all $1 \le s \le k+1$*. Here the constant* C *depends on the upper bound of compliance tensor* $\mathcal{A}$ *but it is independent of the mesh size* h*.*

This result shows that the numerical errors for both unknowns $(\boldsymbol{u}, \underline{\boldsymbol{\sigma}})$ are optimal. In addition, since the only globally-coupled unknown, $\widehat{\boldsymbol{u}}_h$, stays in $\boldsymbol{P}_k(\mathcal{E}_h)$, the order of convergence for the displacement remains optimal only because of a key superconvergence property; see the remark right after Corollary 4.2. In addition, we restrict our result on quasi-uniform meshes to make the proof simple and clear. This result holds for shape-regular meshes also.

2.3. Numerical approximation for nearly incompressible materials.

Here, we consider the numerical approximation of stress for isotropic nearly incompressible materials.

We define isotropic materials to be those whose compliance tensor satisfies the following Assumption 2.1.

Assumption 2.1.

$$\mathcal{A}\underline{\boldsymbol{\tau}} = P_D \underline{\boldsymbol{\tau}}_D + P_T \frac{tr(\underline{\boldsymbol{\tau}})}{3} I_3, \quad where \quad \underline{\boldsymbol{\tau}}_D = \underline{\boldsymbol{\tau}} - \frac{tr(\underline{\boldsymbol{\tau}})}{3} I_3, \tag{2.7}$$

for any $\underline{\boldsymbol{\tau}}$ in $\mathbb{R}^{3\times 3}$, and P_D and P_T are two positive constants. An isotropic material is nearly incompressible if P_T is close to zero.

Theorem 2.2. *If the material is isotropic (whose compliance tensor satisfies Assumption 2.1), P_T is positive, the boundary data $\boldsymbol{g} = 0$, the meshes are quasi-uniform and $\tau = \mathcal{O}(\frac{1}{h})$, then we have*

$$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{L^2(\Omega)} \le C h^s (\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}), \tag{2.8}$$

for all $1 \le s \le k+1$. Here, the constant C is independent of P_T^{-1}.

This result shows that the HDG method (2.2) is locking-free for nearly incompressible materials. We emphasize that the convergence rate of stress for nearly incompressible materials is one order higher than [5, 26] with the same finite element space for numerical trace of displacement.

3. A characterization of the HDG method

In this section we show how to eliminate elementwise the unknowns $\underline{\boldsymbol{\sigma}}_h$ and $\boldsymbol{u}_h$ from the equations (2.2) and rewrite the original system solely in terms of the unknown $\widehat{\boldsymbol{u}}_h$; see also [35]. Via this elimination, we do not have to deal with the large indefinite linear system generated by (2.2), but with the inversion of a sparser symmetric positive definite matrix of remarkably smaller size.

3.1. The local problems.

The result on the above mentioned elimination can be described using additional "local" operators defined as follows:

On each element K, for any $\boldsymbol{\lambda} \in \boldsymbol{M}_h|_{\partial K}$, we denote $(\underline{\boldsymbol{Q}}\boldsymbol{\lambda}, \boldsymbol{U}\boldsymbol{\lambda}) \in \underline{\boldsymbol{V}}(K) \times \boldsymbol{W}(K)$ to be the unique solution of the local problem:

$$(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{\lambda}, \underline{\boldsymbol{v}})_K + (\boldsymbol{U}\boldsymbol{\lambda}, \nabla\cdot\underline{\boldsymbol{v}})_K = \langle \boldsymbol{\lambda}, \underline{\boldsymbol{v}}\cdot\boldsymbol{n}\rangle_{\partial K}, \tag{3.1a}$$

$$-(\nabla\cdot\underline{\boldsymbol{Q}}\boldsymbol{\lambda}, \boldsymbol{\omega})_K + \langle \tau \boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda}, \boldsymbol{\omega}\rangle_{\partial K} = \langle \tau\boldsymbol{\lambda}, \boldsymbol{\omega}\rangle_{\partial K}, \tag{3.1b}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}) \in \underline{\boldsymbol{V}}(K) \times \boldsymbol{W}(K)$.

On each element K, we also denote $(\underline{\boldsymbol{Q}}_S\boldsymbol{\lambda}, \boldsymbol{U}_S\boldsymbol{\lambda}) \in \underline{\boldsymbol{V}}(K) \times \boldsymbol{W}(K)$ to be the unique solution of the local problem:

$$(\mathcal{A}\underline{\boldsymbol{Q}}_S\boldsymbol{f}, \underline{\boldsymbol{v}})_K + (\boldsymbol{U}_S\boldsymbol{f}, \nabla\cdot\underline{\boldsymbol{v}})_K = 0, \tag{3.2a}$$

$$-(\nabla\cdot\underline{\boldsymbol{Q}}_S\boldsymbol{f}, \boldsymbol{\omega})_K + \langle\tau\boldsymbol{P_M}\boldsymbol{U}_S\boldsymbol{f}, \boldsymbol{\omega}\rangle_{\partial K} = -(\boldsymbol{f}, \boldsymbol{\omega})_K, \tag{3.2b}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}) \in \underline{\boldsymbol{V}}(K) \times \boldsymbol{W}(K)$.

It is easy to show the two local problems are well-posed. In addition, due to the linearity of the global system (2.2), the numerical solution $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ satisfies

$$\underline{\boldsymbol{\sigma}}_h = \underline{\boldsymbol{Q}}\widehat{\boldsymbol{u}}_h + \underline{\boldsymbol{Q}}_S\boldsymbol{f}, \quad \boldsymbol{u}_h = \boldsymbol{U}\widehat{\boldsymbol{u}}_h + \boldsymbol{U}_S\boldsymbol{f}. \tag{3.3}$$

3.2. **The global problem.** For the sake of simplicity, we assume the boundary data $\boldsymbol{g} = 0$. Then, the HDG method (2.2) is to find $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h^0$ satisfying

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}_h\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle\widehat{\boldsymbol{u}}_h\,, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.4a}$$

$$-(\nabla\cdot\underline{\boldsymbol{\sigma}}_h\,, \boldsymbol{\omega})_{\mathcal{T}_h} + \langle\tau(\boldsymbol{P_M}\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,, \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = -(\boldsymbol{f}\,, \boldsymbol{\omega})_{\mathcal{T}_h}, \tag{3.4b}$$

$$\langle\underline{\boldsymbol{\sigma}}_h\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} = 0, \tag{3.4c}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h^0$, where $\boldsymbol{M}_h^0 = \{\boldsymbol{\mu} \in \boldsymbol{M}_h : \boldsymbol{\mu}|_{\partial\Omega} = 0\}$.

Combining (3.4c) with (3.3), we have that for all $\boldsymbol{\mu} \in \boldsymbol{M}_h^0$,

$$\langle(\underline{\boldsymbol{Q}}\widehat{\boldsymbol{u}}_h)\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{U}\widehat{\boldsymbol{u}}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} = \langle(\underline{\boldsymbol{Q}}_S\boldsymbol{f})\boldsymbol{n} - \tau\boldsymbol{P_M}\boldsymbol{U}_S\boldsymbol{f}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}. \tag{3.5}$$

Up to now we can see that we only need to solve the reduced global linear system (3.5) first, then recover $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h)$ by (3.3) element by element. Next we show that the global system (3.5) is in fact symmetric positive definite.

3.3. **A characterization of the approximate solution.** The above results suggest the following characterization of the numerical solution of the HDG method.

Theorem 3.1. *The numerical solution of the HDG method* (2.2) *satisfies*

$$\underline{\boldsymbol{\sigma}}_h = \underline{\boldsymbol{Q}}\widehat{\boldsymbol{u}}_h + \underline{\boldsymbol{Q}}_S\boldsymbol{f}, \quad \boldsymbol{u}_h = \boldsymbol{U}\widehat{\boldsymbol{u}}_h + \boldsymbol{U}_S\boldsymbol{f}.$$

If we assume the boundary data $\boldsymbol{g} = 0$, then $\widehat{\boldsymbol{u}}_h \in \boldsymbol{M}_h^0$ is the solution of

$$a_h(\widehat{\boldsymbol{u}}_h, \boldsymbol{\mu}) = \langle(\underline{\boldsymbol{Q}}_S\boldsymbol{f})\boldsymbol{n} - \tau\boldsymbol{P_M}\boldsymbol{U}_S\boldsymbol{f}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}, \quad \forall\boldsymbol{\mu} \in \boldsymbol{M}_h^0, \tag{3.6}$$

where

$$a_h(\widehat{\boldsymbol{u}}_h, \boldsymbol{\mu}) = (\mathcal{A}\underline{\boldsymbol{Q}}\widehat{\boldsymbol{u}}_h, \underline{\boldsymbol{Q}}\boldsymbol{\mu})_{\mathcal{T}_h} + \langle\tau(\boldsymbol{P_M}\boldsymbol{U}\widehat{\boldsymbol{u}}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\mu} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}.$$

In addition, the bilinear operator $a_h(\boldsymbol{\lambda}, \boldsymbol{\lambda})$ is positive definite.

Proof. In order to show (3.6) is true, we only need to show that for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \boldsymbol{M}_h^0$, then

$$a_h(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \langle(\underline{\boldsymbol{Q}}\boldsymbol{\lambda})\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}.$$

According to (3.1), we have

$$(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{m}, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{U}\boldsymbol{m}, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} = \langle \boldsymbol{m}, \underline{\boldsymbol{v}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \tag{3.7a}$$

$$(\nabla\cdot\underline{\boldsymbol{Q}}\boldsymbol{m}, \boldsymbol{\omega})_{\mathcal{T}_h} = \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{m} - \boldsymbol{m}), \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h}, \tag{3.7b}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h$, $\boldsymbol{m} \in \boldsymbol{M}_h^0$. Then, we have

$$\begin{aligned}
&\langle (\underline{\boldsymbol{Q}}\boldsymbol{\lambda})\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \\
=&\langle \boldsymbol{\mu}, (\underline{\boldsymbol{Q}}\boldsymbol{\lambda})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \\
=&(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{\mu}, \underline{\boldsymbol{Q}}\boldsymbol{\lambda})_{\mathcal{T}_h} + (\boldsymbol{U}\boldsymbol{\mu}, \nabla\cdot\underline{\boldsymbol{Q}}\boldsymbol{\lambda})_{\mathcal{T}_h} - \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \quad \text{by (3.7a)} \\
=&(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{\mu}, \underline{\boldsymbol{Q}}\boldsymbol{\lambda})_{\mathcal{T}_h} + (\nabla\cdot\underline{\boldsymbol{Q}}\boldsymbol{\lambda}, \boldsymbol{U}\boldsymbol{\mu})_{\mathcal{T}_h} - \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \\
=&(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{\mu}, \underline{\boldsymbol{Q}}\boldsymbol{\lambda})_{\mathcal{T}_h} + \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{U}\boldsymbol{\mu} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \quad \text{by (3.7b)} \\
=&(\mathcal{A}\underline{\boldsymbol{Q}}\boldsymbol{\mu}, \underline{\boldsymbol{Q}}\boldsymbol{\lambda})_{\mathcal{T}_h} + \langle \tau(\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}), \boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\mu} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \\
=&a_h(\boldsymbol{\lambda}, \boldsymbol{\mu}).
\end{aligned}$$

So, we can conclude that (3.6) holds. We end the proof by showing the bilinear operator $a_h(\cdot,\cdot)$ is positive definite.

If $a_h(\boldsymbol{\lambda}, \boldsymbol{\lambda}) = 0$ for some $\boldsymbol{\lambda} \in \boldsymbol{M}_h^0$, from the previous result we have

$$\underline{\boldsymbol{Q}}\boldsymbol{\lambda} = 0, \quad \boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}|_{\partial\mathcal{T}_h} = 0.$$

We apply integration by parts on (3.1a), we have

$$\langle \underline{\epsilon}(\boldsymbol{U}\boldsymbol{\lambda}), \underline{\boldsymbol{v}}\rangle_{\partial K} = 0, \quad \forall\, \underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}(K).$$

This implies that $\underline{\epsilon}(\boldsymbol{U}\boldsymbol{\lambda})|_K = 0$ for all $K \in \mathcal{T}_h$. So, for any $K \in \mathcal{T}_h$, there are $\boldsymbol{a}_K, \boldsymbol{b}_K \in \mathbb{R}^3$ such that $\boldsymbol{U}\boldsymbol{\lambda}|_K = \boldsymbol{a}_K \times \boldsymbol{x} + \boldsymbol{b}_K$. Since $k \geq 1$, we have $\boldsymbol{P}_M\boldsymbol{U}\boldsymbol{\lambda} = \boldsymbol{U}\boldsymbol{\lambda}$. Combining this result with the fact that $\boldsymbol{P_M}\boldsymbol{U}\boldsymbol{\lambda} - \boldsymbol{\lambda}|_{\partial\mathcal{T}_h} = 0$ and $\boldsymbol{\lambda}|_{\partial\Omega} = 0$, we can conclude that $\boldsymbol{U}\boldsymbol{\lambda} \in \boldsymbol{C}^0(\Omega)$ and $\boldsymbol{U}\boldsymbol{\lambda}|_{\partial\Omega} = 0$.

Finally, let us consider two adjacent elements K_1, K_2 with the interface $F = \bar{K}_1 \cap \bar{K}_2$. In addition, we assume that on K_i, $\boldsymbol{U}\boldsymbol{\lambda}$ can be expressed as

$$\boldsymbol{U}\boldsymbol{\lambda} = \boldsymbol{a}_i \times \boldsymbol{x} + \boldsymbol{b}_i, \quad i = 1, 2.$$

We claim that $\boldsymbol{a}_1 = \boldsymbol{a}_2$ and $\boldsymbol{b}_1 = \boldsymbol{b}_2$. This fact can be shown by considering the continuity of the function on the interface F. We omit the detailed proof since it only involves elementary linear algebra.

From this result we conclude that there exist $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$ such that $\boldsymbol{U}\boldsymbol{\lambda} = \boldsymbol{a}\times\boldsymbol{x} + \boldsymbol{b}$ in Ω. By the fact that $\boldsymbol{U}\boldsymbol{\lambda}|_{\partial\Omega} = 0$, we can conclude that $\boldsymbol{U}\boldsymbol{\lambda} = 0$, hence $\boldsymbol{\lambda} = 0$. This completes the proof. □

Remark 3.2. In Theorem 3.1, we assume the boundary data $\boldsymbol{g} = 0$. Actually, if $\boldsymbol{g}$ is not zero, we can still obtain the same linear system as a_h in Theorem 3.1 by the same treatment of boundary data in [16].

4. Error analysis

In this section we provide detailed proofs for our a priori error estimates — Theorem 2.1 and Theorem 2.2. We use the elementwise Korn's inequality (Lemma 4.1), which is novel and crucial in error analysis. We use $\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}}, \boldsymbol{\Pi}_{\boldsymbol{W}}$ to denote the standard L^2-orthogonal projection onto $\underline{\boldsymbol{V}}_h, \boldsymbol{W}_h$, respectively. In addition, we denote

$$\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}} = \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h, \quad \boldsymbol{e}_{\boldsymbol{u}} = \boldsymbol{\Pi}_{\boldsymbol{W}}\boldsymbol{u} - \boldsymbol{u}_h, \quad \boldsymbol{e}_{\widehat{\boldsymbol{u}}} = \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u} - \widehat{\boldsymbol{u}}_h.$$

In the analysis, we are going to use the following classical results:

$$\|\boldsymbol{u}-\boldsymbol{\Pi_W u}\|_{\Omega} \le Ch^s\|u\|_{s,\Omega} \qquad 0\le s\le k+2, \tag{4.1a}$$

$$\|\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V}}\underline{\boldsymbol{\sigma}}\|_{\Omega} \le Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega} \qquad 0\le t\le k+1, \tag{4.1b}$$

$$\|\boldsymbol{u}-\boldsymbol{P_M u}\|_{\mathcal{E}_h} \le Ch^{s-\frac{1}{2}}\|\boldsymbol{u}\|_{s,\Omega}, \quad 1\le s\le k+1, \tag{4.1c}$$

$$\|\boldsymbol{u}-\boldsymbol{\Pi_W u}\|_{\partial K} \le Ch^{s-\frac{1}{2}}\|\boldsymbol{u}\|_{s,K}, \quad 1\le s\le k+2, \tag{4.1d}$$

$$\|\underline{\boldsymbol{\sigma}}\boldsymbol{n}-\underline{\boldsymbol{\Pi_V}}\underline{\boldsymbol{\sigma}}\boldsymbol{n}\|_{\partial K} \le Ch^{t-\frac{1}{2}}\|\underline{\boldsymbol{\sigma}}\|_{t,K}, \quad 1\le t\le k+1, \tag{4.1e}$$

$$\|\boldsymbol{v}\|_{\partial K} \le Ch^{-\frac{1}{2}}\|\boldsymbol{v}\|_{K}, \qquad \forall\, \boldsymbol{v}\in \boldsymbol{P}_s(K), \tag{4.1f}$$

$$\|\underline{\boldsymbol{\sigma}}\boldsymbol{n}-\boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\|_{\partial K} \le Ch^{t-\frac{1}{2}}\|\underline{\boldsymbol{\sigma}}\|_{t,K}, \quad 1\le t\le k+1. \tag{4.1g}$$

The above results are due to standard approximation theory of polynomials, trace inequality.

Let $\underline{\boldsymbol{\epsilon}}_h$ denote the discrete symmetric gradient operator, such that for any $K\in\mathcal{T}_h$, $\underline{\boldsymbol{\epsilon}}_h|_K=\underline{\boldsymbol{\epsilon}}|_K$. It is well known (see Theorem 2.2 in [14]) that the *kernel* of the operator $\underline{\boldsymbol{\epsilon}}_h(\cdot)$ is:

$$\ker \underline{\boldsymbol{\epsilon}}_h = \Upsilon_h := \{\boldsymbol{\Lambda}\in \boldsymbol{L}^2(\Omega),\ \boldsymbol{\Lambda}|_K=\underline{\boldsymbol{B}}_K\boldsymbol{x}+\boldsymbol{b}_K,\ \underline{\boldsymbol{B}}_K\in\underline{\boldsymbol{\mathcal{A}}}, \boldsymbol{b}_K\in\mathbb{R}^3, K\in\mathcal{T}_h\}.$$

Here, $\underline{\boldsymbol{\mathcal{A}}}$ denotes the set of all anti-symmetric matrices in $\mathbb{R}^{3\times 3}$.

In the analysis, we need the following elementwise Korn's inequality:

Lemma 4.1. *Let $K\in\mathcal{T}_h$ be a generic element with size h_K and $\Upsilon(K):=\Upsilon_h|_K$. Then for any function $\boldsymbol{v}\in\boldsymbol{W}(K)$, we have*

$$\inf_{\boldsymbol{\Lambda}\in\Upsilon(K)}\|\nabla(\boldsymbol{v}+\boldsymbol{\Lambda})\|_K \le C\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_K,$$

Here C is independent of the size h_K. In addition, if K is a tetrahedron, the above inequality holds for any $\boldsymbol{v}\in\boldsymbol{H}^1(K)$.

Proof. Let $\widehat{K}$ denote the reference tetrahedron element and $\boldsymbol{v}\in\boldsymbol{H}^1(K)$. The mapping from $\widehat{K}$ to K is $\boldsymbol{x}=\underline{\boldsymbol{A}}_K\widehat{\boldsymbol{x}}+\boldsymbol{c}_K$ where $\underline{\boldsymbol{A}}_K$ is a non-singular matrix and $\boldsymbol{c}_K\in\mathbb{R}^3$.

We define $\widehat{\boldsymbol{v}}$, which is the pull back of $\boldsymbol{v}$ on $\widehat{K}$, by

$$\underline{\boldsymbol{A}}_K^{-\top}\widehat{\boldsymbol{v}}(\widehat{\boldsymbol{x}})=\boldsymbol{v}(\boldsymbol{x}) \quad \forall \widehat{\boldsymbol{x}}\in\widehat{K}.$$

So, we have

$$\nabla\boldsymbol{v}(\boldsymbol{x})=\nabla(\underline{\boldsymbol{A}}_K^{-\top}\widehat{\boldsymbol{v}})(\boldsymbol{x})=\underline{\boldsymbol{A}}_K^{-\top}(\nabla\widehat{\boldsymbol{v}})(\boldsymbol{x}).$$

The last equality above is due to the fact that every component of $\underline{\boldsymbol{A}}_K^{-\top}$ is constant. It is easy to see that

$$(\nabla\widehat{\boldsymbol{v}})(\boldsymbol{x})=\widehat{\nabla}\widehat{\boldsymbol{v}}(\widehat{\boldsymbol{x}})\underline{\boldsymbol{A}}_K^{-1}.$$

So, we have

$$\underline{\boldsymbol{A}}_K^{-\top}\widehat{\nabla}\widehat{\boldsymbol{v}}(\widehat{\boldsymbol{x}})\underline{\boldsymbol{A}}_K^{-1}=\nabla\boldsymbol{v}(\boldsymbol{x}).$$

By taking the symmetric part of both sides of the above equation, we have

$$\underline{\boldsymbol{A}}_K^{-\top}\widehat{\underline{\boldsymbol{\epsilon}}}\widehat{\boldsymbol{v}}(\widehat{\boldsymbol{x}})\underline{\boldsymbol{A}}_K^{-1} = \underline{\boldsymbol{\epsilon}}\boldsymbol{v}(\boldsymbol{x}). \tag{4.2}$$

According to Theorem 2.3 in [14], the following inequality holds:

$$\inf_{\widehat{\boldsymbol{\Lambda}}\in\Upsilon(\widehat{K})} \|\widehat{\boldsymbol{v}} + \widehat{\boldsymbol{\Lambda}}\|_{1,\widehat{K}} \le C\|\widehat{\underline{\boldsymbol{\epsilon}}}(\widehat{\boldsymbol{v}})\|_{0,\widehat{K}}.$$

So, there is $\widehat{\boldsymbol{\Lambda}} = \underline{\boldsymbol{B}}_{\widehat{K}}\widehat{\boldsymbol{x}} + \boldsymbol{b}_{\widehat{K}}$ with $\underline{\boldsymbol{B}}_{\widehat{K}} \in \underline{\boldsymbol{\mathcal{A}}}$ and $\boldsymbol{b}_{\widehat{K}} \in \mathbb{R}^3$, such that

$$\|\widehat{\nabla}(\widehat{\boldsymbol{v}} + \widehat{\boldsymbol{\Lambda}})\|_{0,\widehat{K}} \le C\|\widehat{\underline{\boldsymbol{\epsilon}}}(\widehat{\boldsymbol{v}})\|_{0,\widehat{K}}. \tag{4.3}$$

We define

$$\boldsymbol{\Lambda}(\boldsymbol{x}) = \underline{\boldsymbol{A}}_K^{-\top}\widehat{\boldsymbol{\Lambda}}(\widehat{\boldsymbol{x}}) \quad \forall \boldsymbol{x} \in K.$$

It is easy to see that

$$\nabla\boldsymbol{\Lambda} = \underline{\boldsymbol{A}}_K^{-\top}\widehat{\nabla}\widehat{\boldsymbol{\Lambda}}\underline{\boldsymbol{A}}_K^{-1} = \underline{\boldsymbol{A}}_K^{-\top}\underline{\boldsymbol{B}}_{\widehat{K}}\underline{\boldsymbol{A}}_K^{-1} \in \underline{\boldsymbol{\mathcal{A}}}.$$

So, $\boldsymbol{\Lambda} \in \Upsilon(K)$. Then, by standard scaling argument with (4.2), (4.3) and the shape regularity of the meshes, we can conclude that the proof for arbitrary tetrahedron element is complete.

Now, we consider the case of arbitrary shape regular element K, which can be hexahedron, prism or pyramid. Let $\boldsymbol{v} = (v_1, v_2, v_3)^\top \in \boldsymbol{W}_h|_K$. It is well known that for any $1 \le i, j, k \le 3$,

$$\partial_j(\partial_k v_i) = \partial_j(\epsilon_{ik}(\boldsymbol{v})) + \partial_k(\epsilon_{ij}(\boldsymbol{v})) - \partial_i(\epsilon_{jk}(\boldsymbol{v})).$$

Here, $\epsilon_{ik}(\boldsymbol{v}) = (\underline{\boldsymbol{\epsilon}}(\boldsymbol{v}))_{ik}$. Consequently, we have

$$\|\nabla(\partial_j v_i - \partial_i v_j)\|_{0,K} \le C\|\nabla\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_{0,K} \le Ch_K^{-1}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_{0,K}.$$

We define an anti-symmetric matrix $\underline{\boldsymbol{B}}_K$ by

$$(\underline{\boldsymbol{B}}_K)_{ij} = \frac{1}{2|K|}\int_K (\partial_j v_i - \partial_i v_j) d\boldsymbol{x} \quad 1 \le i, j \le 3.$$

We take $\boldsymbol{\Lambda} = \underline{\boldsymbol{B}}_K\boldsymbol{x}$, which is obviously in $\Upsilon(K)$. Then, we have

$$\int_K (\nabla(\boldsymbol{v} - \boldsymbol{\Lambda}) - \underline{\boldsymbol{\epsilon}}(\boldsymbol{v}))\, d\boldsymbol{x} = \int_K (\nabla\boldsymbol{v} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{v}))\, d\boldsymbol{x} - \underline{\boldsymbol{B}}_K\int_K 1\, d\boldsymbol{x} = 0.$$

By the Poincaré inequality, we have

$$\|\nabla(\boldsymbol{v} - \boldsymbol{\Lambda}) - \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_{0,K} \le Ch_K \sum_{1\le i,j\le 3}\|\nabla(\partial_j v_i - \partial_i v_j)\|_{0,K} \le C\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_{0,K}.$$

We immediately have that

$$\|\nabla(\boldsymbol{v} - \boldsymbol{\Lambda})\|_{0,K} \le C\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_{0,K}.$$

This completes the proof. □

Step 1 (The error equation). We first present the error equation for the analysis.

Lemma 4.2. *Let* $(\boldsymbol{u}, \underline{\boldsymbol{\sigma}}), (\boldsymbol{u}_h, \underline{\boldsymbol{\sigma}}_h, \widehat{\boldsymbol{u}}_h)$ *solve* (1.1) *and* (2.2), *respectively, then we have*

$$(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{e}_{\boldsymbol{u}}\,,\, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\mathcal{A}(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h}, \tag{4.4a}$$

$$(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = 0, \tag{4.4b}$$

$$\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} = 0, \tag{4.4c}$$

$$\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega} = 0, \tag{4.4d}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h$.

Proof. We notice that the exact solution $(\boldsymbol{u}, \underline{\boldsymbol{\sigma}}, \boldsymbol{u}|_{\mathcal{E}_h})$ also satisfies the equation (2.2). Hence, after simple algebraic manipulations, we get that

$$\begin{aligned}
&(\mathcal{A}\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{\Pi}_{\boldsymbol{W}}\boldsymbol{u}\,,\, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\qquad = -(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + \langle \boldsymbol{u} - \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{u} - \boldsymbol{\Pi}_{\boldsymbol{W}}\boldsymbol{u}\,,\, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h},\\
&(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}\,,\, \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = -(\boldsymbol{f}\,,\, \boldsymbol{\omega})_{\mathcal{T}_h} - (\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h}\\
&\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}\,,\, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} = 0,\\
&\langle \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega} = -\langle \boldsymbol{u} - \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega},
\end{aligned}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{w}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h$. Notice that the local spaces satisfy the following inclusion property:

$$\nabla\cdot\underline{\boldsymbol{V}}(K) \subset \boldsymbol{W}(K), \quad \underline{\boldsymbol{\epsilon}}(\boldsymbol{W}(K)) \subset \underline{\boldsymbol{V}}(K), \quad \underline{\boldsymbol{V}}(K)\boldsymbol{n}|_F \subset \boldsymbol{M}(F).$$

Hence by the property of the L^2-projection, the above system can be simplified as:

$$\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{\Pi}_{\boldsymbol{W}}\boldsymbol{u}\,,\, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= -(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h},\\
(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}\,,\, \boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} &= -(\boldsymbol{f}\,,\, \boldsymbol{\omega})_{\mathcal{T}_h},\\
\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}\,,\, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} &= 0,\\
\langle \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{u}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega} &= 0,
\end{aligned}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{w}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \boldsymbol{M}_h$. Here we applied the fact that $(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \nabla\boldsymbol{\omega})_{\mathcal{T}_h} = (\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\, \underline{\boldsymbol{\epsilon}}(\boldsymbol{\omega}))_{\mathcal{T}_h} = 0$. If we now subtract the equations (2.2), we obtain the result. This completes the proof. $\square$

Step 2 (Estimate of $\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}$). We are now ready to obtain our first estimate.

Proposition 4.1. *We have*

$$(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + \langle \tau(\boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\,,\, \boldsymbol{P}_{\boldsymbol{M}}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} = -(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + T_1 - T_2,$$

where T_1, T_2 *are defined as:*

$$\begin{aligned}
T_1 &:= \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - (\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\
T_2 &:= \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \tau(\boldsymbol{P}_{\boldsymbol{M}}(\boldsymbol{u} - \boldsymbol{\Pi}_{\boldsymbol{W}}\boldsymbol{u}))\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

Proof. By the error equation (4.4d) we know that $\boldsymbol{e}_{\widehat{\boldsymbol{u}}} = 0$ on $\partial\Omega$. This implies that

$$\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\rangle_{\partial\Omega} = 0.$$

Now taking $(\underline{\boldsymbol{v}}, \boldsymbol{w}, \boldsymbol{\mu}) = (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}, \boldsymbol{e}_{\boldsymbol{u}}, \boldsymbol{e}_{\widehat{\boldsymbol{u}}})$ in error equations (4.4a)–(4.4c) and adding these equations together with the above identity, we obtain, after some algebraic manipulation,

$$(4.5)\quad (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n} - (\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} = -(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h}.$$

Now we work with the second term on the left-hand side,

$$\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n} - (\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}) = \underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\boldsymbol{\sigma}}_h\boldsymbol{n} - \underline{\boldsymbol{\sigma}}\boldsymbol{n} + \widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}$$

by the definition of the numerical trace (2.2e),

$$\begin{aligned} &= -(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\boldsymbol{n} - \tau(\boldsymbol{P_M}\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h),\\ &= -(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\boldsymbol{n} + \tau(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}) - \tau(\boldsymbol{P_M}(\boldsymbol{\Pi_W}\boldsymbol{u} - \boldsymbol{u})),\end{aligned}$$

the last step is by the definition of $\boldsymbol{e}_{\boldsymbol{u}}, \boldsymbol{e}_{\widehat{\boldsymbol{u}}}$. Inserting the above identity into (4.5) and moving terms around, we have

$$\begin{aligned}(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} &+ \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\rangle_{\partial\mathcal{T}_h}\\ &= -(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + T_1 - T_2.\end{aligned}$$

Finally, notice that on each $F \in \partial\mathcal{T}_h$, $\tau(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})|_F \in \boldsymbol{M}(F)$, so we have

$$\langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\rangle_{\partial\mathcal{T}_h} = \langle \boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\rangle_{\partial\mathcal{T}_h}.$$

This completes the proof. □

From the above energy argument we can see that we need to bound T_1, T_2 in order to have an estimate for $\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}$. Next we present the estimates for these two terms:

Lemma 4.3. *If the parameter $\tau = \mathcal{O}(h^{-1})$, we have*

$$\begin{aligned} T_1 &\le Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\,(\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h} + \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e}_{\boldsymbol{u}})\|_{\mathcal{T}_h}),\\ T_2 &\le Ch^{s-1}\|\boldsymbol{u}\|_{s,\Omega}\,\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h},\end{aligned}$$

for all $1 \le t \le k+1, 1 \le s \le k+2$.

Proof. We first bound T_2. We have

$$\begin{aligned} T_2 &= \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{P_M}(\boldsymbol{u} - \boldsymbol{\Pi_W}\boldsymbol{u}))\rangle_{\partial\mathcal{T}_h} = \langle \boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{P_M}(\boldsymbol{u} - \boldsymbol{\Pi_W}\boldsymbol{u}))\rangle_{\partial\mathcal{T}_h}\\ &= \langle \boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\tau(\boldsymbol{u} - \boldsymbol{\Pi_W}\boldsymbol{u})\rangle_{\partial\mathcal{T}_h}\\ &\le \|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h}\tau^{\frac{1}{2}}\|\boldsymbol{u} - \boldsymbol{\Pi_W}\boldsymbol{u}\|_{\partial\mathcal{T}_h}\\ &\le Ch^s(\tau^{\frac{1}{2}}h^{-\frac{1}{2}})\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{u}\|_{s,\Omega},\end{aligned}$$

for all $1 \le s \le k+2$. In the last step we applied the inequality (4.1d).

The estimate for T_1 is much more sophisticated. We first split T_1 into two parts:

$$T_1 = T_{11} + T_{12},$$

where

$$\begin{aligned} T_{11} &:= \langle \boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{\sigma}}\boldsymbol{n} - (\underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\ T_{12} &:= \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{P_M}\boldsymbol{e}_{\boldsymbol{u}}\,,\,\underline{\boldsymbol{\sigma}}\boldsymbol{n} - (\underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

For T_{11}, we simply apply the Cauchy-Schwarz inequality,

$$\begin{aligned} T_{11} &\le \|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h}\, \tau^{-\frac{1}{2}}\|\underline{\boldsymbol{\sigma}}\boldsymbol{n} - (\underline{\boldsymbol{\Pi_V \sigma}})\boldsymbol{n}\|_{\partial\mathcal{T}_h} \\ &\le Ch^t(\tau^{-\frac{1}{2}}h^{-\frac{1}{2}})\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h}, \end{aligned}$$

for all $1 \le t \le k+1$. Here we used the inequality (4.1e).

Now we work on T_{12}. Using the L^2-orthogonal property of the projection $\boldsymbol{P_M}$, we can write

$$\begin{aligned} T_{12} &= \langle \boldsymbol{e_u} - \boldsymbol{P_M e_u}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - (\underline{\boldsymbol{\Pi_V \sigma}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ &= \langle \boldsymbol{e_u} - \boldsymbol{P_M e_u}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ &= \langle \boldsymbol{e_u} - \boldsymbol{P_M e_u}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial\mathcal{T}_h}, \end{aligned}$$

by the fact that $\underline{\boldsymbol{\Pi_V \sigma}}\boldsymbol{n}|_F, \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})|_F \in \boldsymbol{M}(F)$ for all $F \in \partial\mathcal{T}_h$,

$$\begin{aligned} T_{12} &= \langle \boldsymbol{e_u}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial\mathcal{T}_h}, \quad \text{since } \boldsymbol{P_M e_u}|_F \in \boldsymbol{M}(F), \forall F \in \partial\mathcal{T}_h, \\ &= \langle \boldsymbol{e_u} + \boldsymbol{\Lambda}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial\mathcal{T}_h}, \end{aligned}$$

where $\boldsymbol{\Lambda} \in \boldsymbol{L}^2(\Omega)$ is any vector-valued function in Υ_h. Notice here the last step holds only if $\Upsilon_h|_F \in \boldsymbol{M}(F)$, $\forall F \in \partial\mathcal{T}_h$. This is true if $k \ge 1$. Next, on each $K \in \mathcal{T}_h$, if we denote $\overline{\boldsymbol{u}}$ to be the average of $\boldsymbol{u}$ over K, then we have

$$\begin{aligned} \langle \boldsymbol{e_u} + \boldsymbol{\Lambda}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial K} &= \langle \boldsymbol{e_u} + \boldsymbol{\Lambda} - \overline{(\boldsymbol{e_u} + \boldsymbol{\Lambda})}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial K} \\ &\le \|\boldsymbol{e_u} + \boldsymbol{\Lambda} - \overline{(\boldsymbol{e_u} + \boldsymbol{\Lambda})}\|_{\partial K}\|\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\|_{\partial K}, \end{aligned}$$

by the standard inequalities (4.1f), (4.1g),

$$\begin{aligned} \langle \boldsymbol{e_u} + \boldsymbol{\Lambda}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial K} &\le Ch^{t-1}\|\underline{\boldsymbol{\sigma}}\|_{t,K}\|\boldsymbol{e_u} + \boldsymbol{\Lambda} - \overline{(\boldsymbol{e_u} + \boldsymbol{\Lambda})}\|_K \\ &\le Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,K}\|\nabla(\boldsymbol{e_u} + \boldsymbol{\Lambda})\|_K, \end{aligned}$$

for all $1 \le t \le k+1$. The last step is by the Poincaré inequality. Notice that the constant C in the above inequality is independent of $\boldsymbol{\Lambda} \in \Upsilon_h$. Now applying the Lemma 4.1, yields,

$$\langle \boldsymbol{e_u} + \boldsymbol{\Lambda}\,,\, \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\rangle_{\partial K} \le Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,K}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_K.$$

Summing over all $K \in \mathcal{T}_h$, we have

$$T_{12} \le Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h},$$

for all $1 \le t \le k+1$. We complete the proof by combining the estimates for T_2, T_{11}, T_{12}. □

Combining Lemma 4.3 and Proposition 4.1, we obtain the following estimate.

Corollary 4.1. *If the parameter $\tau = \mathcal{O}(h^{-1})$, then we have*

$$\begin{aligned} &\|\underline{\boldsymbol{e_\sigma}}\|^2_{L^2(\mathcal{A},\Omega)} + \|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|^2_{\partial\mathcal{T}_h} \\ &\qquad \le C\left(h^{2t}\|\underline{\boldsymbol{\sigma}}\|^2_{t,\Omega} + h^{2(s-1)}\|\boldsymbol{u}\|^2_{s,\Omega} + h^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h}\right), \end{aligned}$$

for all $1 \le s \le k+2, 1 \le t \le k+1$, the constant C is independent of h and the exact solution.

The proof is omitted. One can obtain the above result by the Cauchy-Schwarz inequality and the weighted Young's inequality. Finally, we can finish the estimate for $\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}$ by the following estimate for $\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})$:

Lemma 4.4. *Under the same assumption as Theorem* 4.1*, we have*

$$\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} \le C\left(h^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega} + \|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)} + \|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h}\right),$$

for all $0 \le t \le k+1$.

Proof. Notice that $\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}) \in \underline{\boldsymbol{V}}_h$, so we can take $\underline{\boldsymbol{v}} = \underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})$ in the error equation (4.4a), after integrating by parts, we have:

$$\begin{aligned}(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h} - (\nabla\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h} + \langle\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ = (\mathcal{A}(\underline{\boldsymbol{\Pi_V\sigma}} - \underline{\boldsymbol{\sigma}})\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h}.\end{aligned}$$

Notice that $\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}) \in \underline{\boldsymbol{V}}_h$ and it is symmetric, so we have

$$(\nabla\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h} = \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|^2_{\mathcal{T}_h}, \quad \langle\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

Inserting these two identities into the first equation, we have

$$\begin{aligned}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|^2_{\mathcal{T}_h} &= (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h} + \langle\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &\quad + (\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V\sigma}})\,,\,\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u}))_{\mathcal{T}_h}\\ &\le C\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} + C\tau^{-\frac{1}{2}}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\boldsymbol{n}\|_{\partial\mathcal{T}_h}\\ &\quad + Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h}\\ &\le C\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} + C\tau^{-\frac{1}{2}}h^{-\frac{1}{2}}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h}\\ &\quad + Ch^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} \qquad \text{by inverse inequality (4.1f).}\end{aligned}$$

The proof is complete by the assumption $\tau = \mathcal{O}(h^{-1})$. □

Finally, combining Lemma 4.4 and Theorem 4.1, after simple algebraic manipulation, we have our first error estimate:

Corollary 4.2. *Under the same assumption as in Theorem* 4.1*, we have*

$$\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)} + \|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h} + \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} \le C(h^t\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega} + h^{s-1}\|\boldsymbol{u}\|_{s,\Omega}),$$

for all $1 \le t \le k+1, 1 \le s \le k+2$*, the constant* C *is independent of* h *and the exact solution.*

One can see that by taking $t = k+1, s = k+2$, both of the errors $\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}, \underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})$ obtain optimal convergence rate. Moreover, if we take $\tau = 1/h$, we readily obtain the superconvergence property

$$\|h^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h} \le C\,h^{k+2},$$

for smooth solutions. It is this superconvergence property, the one which allows us to obtain the optimal convergence in the stress and, as we are going to see next, in the displacement.

Step 3 (Estimate of $\boldsymbol{e_u}$). Next we use a standard duality argument to get an estimate for $\boldsymbol{e_u}$. First we present an important identity.

Proposition 4.2. *Assume that* $(\boldsymbol{\phi}, \underline{\boldsymbol{\psi}}) \in \boldsymbol{H}^2(\Omega) \times \underline{\boldsymbol{H}}^1(\Omega)$ *is the solution of the adjoint problem* (2.3a)*, we have*

$$\begin{aligned}\|\boldsymbol{e_u}\|_\Omega^2 &= (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} - (\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} \\ &\quad - \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n} - (\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n})\,,\, \boldsymbol{\phi} - \boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{e_u} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Proof. By the dual equation (2.3), we can write

$$\begin{aligned}\|\boldsymbol{e_u}\|_\Omega^2 &= (\boldsymbol{e_u}\,,\, \nabla\cdot\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \mathcal{A}\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{\phi}))_{\mathcal{T}_h} \\ &= (\boldsymbol{e_u}\,,\, \nabla\cdot\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \nabla\boldsymbol{\phi})_{\mathcal{T}_h} \\ &= (\boldsymbol{e_u}\,,\, \nabla\cdot\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \nabla\boldsymbol{\Pi_W\phi})_{\mathcal{T}_h} \\ &\quad + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\boldsymbol{e_u}\,,\, \nabla\cdot(\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}}))_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \nabla(\boldsymbol{\phi} - \boldsymbol{\Pi_W\phi}))_{\mathcal{T}_h},\end{aligned}$$

integrating by parts for the last two terms, applying the property of the L^2-projections, yields

$$\begin{aligned}\|\boldsymbol{e_u}\|_\Omega^2 &= (\boldsymbol{e_u}\,,\, \nabla\cdot\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \nabla\boldsymbol{\Pi_W\phi})_{\mathcal{T}_h} \\ &\quad + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + \langle \boldsymbol{e_u}\,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n}\,,\, \boldsymbol{\phi} - \boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Taking $\underline{\boldsymbol{v}} := \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}}$ and $\boldsymbol{\omega} := \boldsymbol{\Pi_W\phi}$ in the error equations (4.4a) and (4.4b), respectively, and inserting these two equations into the above identity, we obtain

$$\begin{aligned}\|\boldsymbol{e_u}\|_\Omega^2 &= \langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\sigma}})\,,\, \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} - \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h} \\ &\quad + (\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + \langle \boldsymbol{e_u}\,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}\boldsymbol{\psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n}\,,\, \boldsymbol{\phi} - \boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Next, note that by the regularity assumption, $(\underline{\boldsymbol{\psi}}, \boldsymbol{\phi}) \in \underline{\boldsymbol{H}}^2(\Omega) \times \boldsymbol{H}^1(\Omega)$, so the normal component of $\underline{\boldsymbol{\psi}}$ and $\boldsymbol{\phi}$ are continuous across each face $F \in \mathcal{E}_h$. By the equation (2.2c), the normal component of $\underline{\widehat{\boldsymbol{\sigma}}}_h$ is also strongly continuous across each face $F \in \mathcal{E}_h$. This implies that

$$\begin{aligned}-\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= -\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\Omega} = 0, &&\text{by (4.4d)},\\ \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} &= \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\phi}\rangle_{\partial\Omega} = 0 &&\text{by (2.3c)}.\end{aligned}$$

Adding these two zero terms into the previous equation and rearranging the terms, we obtain the expression as presented in the proposition. ☐

As a consequence of the result just proved, we can obtain our estimate of $\boldsymbol{e_u}$.

Corollary 4.3. *Under the same assumption as in Theorem* 4.1*, in addition, if the elliptic regularity property* (2.4) *holds, then we have*

$$\|\boldsymbol{e_u}\|_\Omega \le C(h^{t+1}\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega} + h^s\|\boldsymbol{u}\|_{s,\Omega}),$$

for $1 \le t \le k+1, 1 \le s \le k+2$.

Proof. We will estimate each of the terms on the right-hand side of the identity in Proposition 4.2.

$$(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\psi}}-\underline{\boldsymbol{\Pi_V\psi}})_{\mathcal{T}_h}\leq Ch\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)}\|\underline{\boldsymbol{\psi}}\|_{1,\Omega}\leq Ch\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\mathcal{A},\Omega)}\|\boldsymbol{e_u}\|_\Omega,$$

by the projection property (4.1b) and the regularity assumption (2.4).

$$\begin{aligned}(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\,,\,\underline{\boldsymbol{\Pi_V\psi}})_{\mathcal{T}_h}&=(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\,,\,\underline{\boldsymbol{\psi}})_{\mathcal{T}_h}-(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\,,\,\underline{\boldsymbol{\psi}}-\underline{\boldsymbol{\Pi_V\psi}})_{\mathcal{T}_h}\\&=(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}}\,,\,\mathcal{A}\underline{\boldsymbol{\psi}}-\overline{\mathcal{A}\underline{\boldsymbol{\psi}}})_{\mathcal{T}_h}-(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\,,\,\underline{\boldsymbol{\psi}}-\underline{\boldsymbol{\Pi_V\psi}})_{\mathcal{T}_h}\\&\leq Ch\|\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}}\|_\Omega\|\underline{\boldsymbol{\psi}}\|_{1,\Omega}\\&\leq Ch\|\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}}\|_\Omega\|\boldsymbol{e_u}\|_\Omega\\&\leq Ch^{t+1}\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\boldsymbol{e_u}\|_\Omega,\end{aligned}$$

for all $0\leq t\leq k+1$. Here we applied the Galerkin orthogonal property of the local L^2-projection $\underline{\boldsymbol{\Pi_V}}$ and the regularity assumption (2.4).

For the third term, by the definition of the numerical trace (2.2e), we have

$$\begin{aligned}\langle\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n}-(\underline{\boldsymbol{\sigma}}\boldsymbol{n}-\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}&=-\langle(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&\quad+\langle\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}-\underline{\boldsymbol{\sigma}}_h\boldsymbol{n}\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&=-\langle(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&\quad-\langle\tau(\boldsymbol{P_M}\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h)\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&=-\langle(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&\quad+\langle\tau(\boldsymbol{P_M}\boldsymbol{e_u}-\boldsymbol{e}_{\widehat{\boldsymbol{u}}})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}\\&\quad+\langle\tau(\boldsymbol{P_M}(\boldsymbol{u}-\boldsymbol{\Pi_W u})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

To bound the third term, we only need to bound the above three terms individually.

$$\begin{aligned}\langle(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}&\leq\|(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\|_{\partial\mathcal{T}_h}\\&\leq Ch^{t-\frac{1}{2}}\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}h^{\frac{3}{2}}\|\boldsymbol{\phi}\|_{2,\Omega},\end{aligned}$$

by the standard inequalities , (4.1d), (4.1e),

$$\leq Ch^{t+1}\|\underline{\boldsymbol{\sigma}}\|_{t,\Omega}\|\boldsymbol{e_u}\|_\Omega,$$

for all $1\leq t\leq k+1$. The last step is due to the regularity assumption (2.4).

Similarly, we apply the Cauchy-Schwarz inequality and (4.1d) for the other two terms:

$$\begin{aligned}\langle\tau(\boldsymbol{P_M}\boldsymbol{e_u}-\boldsymbol{e}_{\widehat{\boldsymbol{u}}})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}&\leq C\tau^{\frac{1}{2}}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e_u}-\boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\|_{\partial\mathcal{T}_h}\\&\leq C\tau^{\frac{1}{2}}h^{\frac{3}{2}}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e_u}-\boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\phi}\|_{2,\Omega}\\&\leq C\tau^{\frac{1}{2}}h^{\frac{3}{2}}\|\tau^{\frac{1}{2}}(\boldsymbol{P_M}\boldsymbol{e_u}-\boldsymbol{e}_{\widehat{\boldsymbol{u}}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{e_u}\|_\Omega.\end{aligned}$$

$$\begin{aligned}\langle\tau(\boldsymbol{P_M}(\boldsymbol{u}-\boldsymbol{\Pi_W u})\,,\,\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\rangle_{\partial\mathcal{T}_h}&\leq\tau\|\boldsymbol{P_M}(\boldsymbol{u}-\boldsymbol{\Pi_W u})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\|_{\partial\mathcal{T}_h}\\&\leq\tau\|\boldsymbol{u}-\boldsymbol{\Pi_W u}\|_{\partial\mathcal{T}_h}\|\boldsymbol{\phi}-\boldsymbol{\Pi_W\phi}\|_{\partial\mathcal{T}_h}\\&\leq C\tau h^{s-\frac{1}{2}}\|\boldsymbol{u}\|_{s,\Omega}h^{\frac{3}{2}}\|\boldsymbol{\phi}\|_{2,\Omega}\\&\leq C\tau h^{s+1}\|\boldsymbol{u}\|_{s,\Omega}\|\boldsymbol{e_u}\|_\Omega,\end{aligned}$$

for all $1\leq s\leq k+2$.

Finally, for the last term in Proposition 4.2, we can write:

$$\begin{aligned}\langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi_V \psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} =& \langle \boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi_V \psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ &+ \langle \boldsymbol{e_u} - \boldsymbol{P_M e_u} \,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi_V \psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

For the first term, we can apply a similar argument as in the previous steps to obtain:

$$\langle \boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi_V \psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \le C\tau^{-\frac{1}{2}} h^{\frac{1}{2}} \|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h} \|\boldsymbol{e_u}\|_{\Omega}.$$

For the second term, we apply the same argument for the estimate of T_{12} in the proof of Lemma 4.3 and obtain:

$$\begin{aligned}\langle \boldsymbol{e_u} - \boldsymbol{P_M e_u} \,,\, (\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi_V \psi}})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &\le Ch\|\underline{\boldsymbol{\psi}}\|_{1,\Omega}\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h} \\ &\le Ch\|\underline{\boldsymbol{\epsilon}}(\boldsymbol{e_u})\|_{\mathcal{T}_h}\|\boldsymbol{e_u}\|_{\Omega}.\end{aligned}$$

Finally, if we take $\tau = \mathcal{O}(h^{-1})$ and combine all the above estimates and Proposition 4.2, we obtain the estimate in Lemma 4.3. $\square$

As a consequence of Proposition 4.2, Lemma 4.3, we can obtain Theorem 2.1 by a simple triangle inequality and the approximation property of the projections $\boldsymbol{\Pi_W}, \underline{\boldsymbol{\Pi_V}}$ (4.1a), (4.1b).

Step 4 (Proof of locking-free result). We can now give the proof of Theorem 2.2.

Proof of Theorem 2.2. In what follows, we assume that s is some arbitrary real number in $[1, k+1]$ and C is a positive constant independent of P_T and s. We recall that $\underline{\boldsymbol{e_\sigma}} = \underline{\boldsymbol{\Pi_V \sigma}} - \underline{\boldsymbol{\sigma}}_h, \boldsymbol{e_u} = \boldsymbol{\Pi_W u} - \boldsymbol{u}_h, \boldsymbol{e_{\widehat{u}}} = \boldsymbol{P_M u} - \widehat{\boldsymbol{u}}_h$.

For any $B \in \mathbb{R}^{3\times 3}$, we denote $B^D := B - \frac{1}{3}\mathrm{tr}B\, I_3$. So, we have

$$\underline{\boldsymbol{e_\sigma}} = \underline{\boldsymbol{e_\sigma}}^D + \frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}} I_3.$$

By Assumption 2.1 and Theorem 2.1, we have

$$\|P_D^{\frac{1}{2}}\underline{\boldsymbol{e_\sigma}}^D\|_{L^2(\Omega)} \le \|\underline{\boldsymbol{e_\sigma}}\|_{L^2(\mathcal{A},\Omega)} \le Ch^s(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}). \tag{4.6}$$

In order to bound $\|\mathrm{tr}\underline{\boldsymbol{e_\sigma}}\|_{L^2(\Omega)}$ independently of P_T^{-1}, we would like to use the well-known result [12, 13] that for any $q \in L^2(\Omega)$ with $\int_\Omega q dx = 0$, we have

$$\|q\|_{L^2(\Omega)} \le C_0 \sup_{\boldsymbol{\eta}\in \boldsymbol{H}_0^1(\Omega)} \frac{(q, \nabla\cdot\boldsymbol{\eta})_\Omega}{\|\boldsymbol{\eta}\|_{H^1(\Omega)}}, \tag{4.7}$$

for C_0 solely depends on the domain Ω. By the assumption $\boldsymbol{g} = 0$, taking $\underline{\boldsymbol{v}} = I_3$ in (4.4a), we have that $\int_\Omega \mathrm{tr}(\mathcal{A}\underline{\boldsymbol{e_\sigma}})dx = 0$. According to Assumption 2.1 and the fact that $P_T > 0$, we have $\int_\Omega \mathrm{tr}\underline{\boldsymbol{e_\sigma}} dx = 0$.

For any $\boldsymbol{\eta} \in \boldsymbol{H}_0^1(\Omega)$, we have

$$\begin{aligned}(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}}, \nabla\cdot\boldsymbol{\eta})_\Omega =& -(\nabla(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}}), \boldsymbol{\eta})_{\mathcal{T}_h} + \langle(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}})\boldsymbol{n}, \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\ =& -(\nabla(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}}), \boldsymbol{\Pi_W\eta})_{\mathcal{T}_h} + \langle(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}})\boldsymbol{n}, \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\ =&(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}}, \nabla\cdot\boldsymbol{\Pi_W\eta})_{\mathcal{T}_h} - \langle(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}})\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\ =&(\underline{\boldsymbol{e_\sigma}} - \underline{\boldsymbol{e_\sigma}}^D, \nabla\boldsymbol{\Pi_W\eta})_{\mathcal{T}_h} - \langle(\underline{\boldsymbol{e_\sigma}} - \underline{\boldsymbol{e_\sigma}}^D)\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

By (4.4b) with $\boldsymbol{\omega} = \boldsymbol{\Pi_W \eta}$, we have

$$\begin{aligned}
&(\frac{1}{3}\mathrm{tr}\underline{\boldsymbol{e_\sigma}}, \nabla\cdot\boldsymbol{\eta})_\Omega \\
=&\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta}\rangle_{\partial\mathcal{T}_h} - (\underline{\boldsymbol{e_\sigma}}^D, \nabla\boldsymbol{\Pi_W\eta})_{\mathcal{T}_h} - \langle(\underline{\boldsymbol{e_\sigma}} - \underline{\boldsymbol{e_\sigma}}^D)\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\
=&T_1 + T_2,
\end{aligned}$$

where

$$\begin{aligned}
T_1 :=&\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e_\sigma}}\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}, \\
T_2 :=& - (\underline{\boldsymbol{e_\sigma}}^D, \nabla\boldsymbol{\Pi_W\eta})_{\mathcal{T}_h} + \langle \underline{\boldsymbol{e_\sigma}}^D\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

For the bound of T_1, by (2.2c) and the fact that $\boldsymbol{\eta} = 0$ on $\partial\Omega$, we have

$$\begin{aligned}
&\langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta}\rangle_{\partial\mathcal{T}_h} \\
&= \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta} - \boldsymbol{P_M\eta}\rangle_{\partial\mathcal{T}_h} \\
&= \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\
&= \langle (\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} + \langle \underline{\boldsymbol{e_\sigma}}\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\
&\quad + \langle \tau(\boldsymbol{P_M u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

So, we have

$$\begin{aligned}
T_1 = &\langle (\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi_V\sigma}})\boldsymbol{n}, \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h} \\
&+ \langle \tau(\boldsymbol{P_M u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{\Pi_W\eta} - \boldsymbol{\eta}\rangle_{\partial\mathcal{T}_h}.
\end{aligned} \tag{4.8}$$

According to Corollary 4.2, we have

$$\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|_{\partial\mathcal{T}_h} \leq Ch^s(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}). \tag{4.9}$$

By the definition of $\boldsymbol{e_{\widehat{u}}}$ and $\boldsymbol{e_u}$, we have

$$\begin{aligned}
\|\tau^{\frac{1}{2}}(\boldsymbol{P_M u}_h - \widehat{\boldsymbol{u}}_h)\|^2_{\partial\mathcal{T}_h} &\leq 2\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|^2_{\partial\mathcal{T}_h} + 2\|\tau^{\frac{1}{2}}\boldsymbol{P_M}(\boldsymbol{u} - \boldsymbol{\Pi_W u})\|^2_{\partial\mathcal{T}_h} \\
&\leq 2\|\tau^{\frac{1}{2}}(\boldsymbol{P_M e_u} - \boldsymbol{e_{\widehat{u}}})\|^2_{\partial\mathcal{T}_h} + 2\|\tau^{\frac{1}{2}}(\boldsymbol{u} - \boldsymbol{\Pi_W u})\|^2_{\partial\mathcal{T}_h}.
\end{aligned}$$

Now applying Young's inequality and (4.9), (4.1a) we obtain:

$$\|\tau^{\frac{1}{2}}(\boldsymbol{P_M u}_h - \widehat{\boldsymbol{u}}_h)\|_{\partial\mathcal{T}_h} \leq Ch^s(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}). \tag{4.10}$$

According to (4.8), (4.10), we have

$$T_1 \leq Ch^s(\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega})\|\boldsymbol{\eta}\|_{H^1(\Omega)}. \tag{4.11}$$

For the bound of T_2, we have

$$\begin{aligned} T_2 =& -(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \nabla \boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{\eta})_{\mathcal{T}_h} + \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D \boldsymbol{n}, \boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{\eta} - \boldsymbol{\eta} \rangle_{\partial \mathcal{T}_h} \\ =& -(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \nabla (\boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{\eta} - \boldsymbol{\eta}))_{\mathcal{T}_h} + \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D \boldsymbol{n}, \boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{\eta} - \boldsymbol{\eta} \rangle_{\partial \mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \nabla \boldsymbol{\eta})_{\mathcal{T}_h} \\ =& (\nabla \cdot \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{\eta} - \boldsymbol{\eta})_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \nabla \boldsymbol{\eta})_{\mathcal{T}_h} \\ =& -(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}^D, \nabla \boldsymbol{\eta})_{\mathcal{T}_h}. \end{aligned}$$

By (4.6), we have

$$T_2 \leq C h^s (\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}) \|\boldsymbol{\eta}\|_{H^1(\Omega)}. \tag{4.12}$$

Finally, combining the estimates (4.6), (4.7), (4.11), (4.12), we have

$$\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{L^2(\Omega)} \leq C_1 h^s (\|\boldsymbol{u}\|_{s+1,\Omega} + \|\underline{\boldsymbol{\sigma}}\|_{s,\Omega}) \|\boldsymbol{\eta}\|_{H^1(\Omega)}.$$

Here the constant C_1 is independent of P_T^{-1}. □

5. Numerical experiment

In this section, we display numerical experiments in 2D to verify the error estimates provided in Theorem 2.1. We also display numerical results showing that our method does not exhibit volumetric-locking when the material tends to be incompressible. In addition, our numerical results suggest that the error estimates provided in Theorem 2.2 for the incompressible limit case are *sharp*.

We carry out the numerical experiments on the domain $\Omega = (0,1) \times (0,1)$ and monitor the errors $\|\underline{\boldsymbol{\Pi}_{\boldsymbol{V}} \boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{L^2(\Omega)}$ and $\|\boldsymbol{\Pi}_{\boldsymbol{W}} \boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$.To explore the dependence of the convergence properties of our method with respect to the form of the meshes, we consider two types of meshes, as shown in Figure 1.

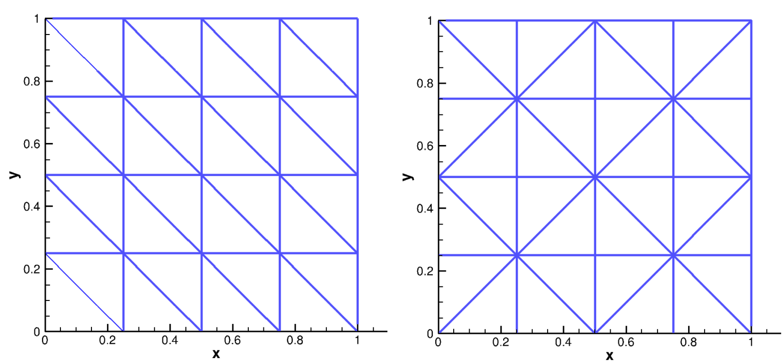

FIGURE 1. An example of Mesh-1(left) and Mesh-2(right) with $h = 0.354$

5.1. **Order of convergence of our HDG method.**

In this section, we consider an isotropic material in 2D with plain stress condition and take the Poisson Ratio $\nu = 0.3$ and the Young's Modulus $E = 1$:

$$\mathcal{A} \underline{\boldsymbol{\sigma}} = \frac{1+\nu}{E} \underline{\boldsymbol{\sigma}} - \frac{\nu}{E} \operatorname{tr}(\underline{\boldsymbol{\sigma}}) I_2. \tag{5.1}$$

In particular, we test our HDG method on a smooth solution $\boldsymbol{u} = (u_1, u_2)$ in [35], such that:

$$u_1 = 10 \sin(\pi x)(1-x)(y-y^2)(1-0.5y), \qquad u_2 = 0. \tag{5.2}$$

We set $\boldsymbol{f}$ and $\boldsymbol{g}$ to satisfy the above exact solution (5.2). To explore the convergence properties of our method, we conduct numerical experiments for $k = 0, 1, 2, 3$ and take $\tau = \mathcal{O}(\frac{1}{h})$. The history of convergence is displayed in Table 2. We observe that when $k \geq 1$, our method converges with order $k+1$ in the stress and order $k+2$ in the displacement for both Mesh-1 and Mesh-2. In addition, the numerical results suggest that our method does not converge to the exact solution when $k = 0$. To aid visualization, we also plot the convergence sequence of the displacement in Figure 2.

TABLE 2. History of convergence for the exact solution (5.2) where $h = 0.177$.

		$\Vert\mathbf{\Pi}_{\boldsymbol{V}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\Vert_{L^2(\Omega)}$		$\Vert\mathbf{\Pi}_{\boldsymbol{W}}\boldsymbol{u} - \boldsymbol{u}_h\Vert_{L^2(\Omega)}$		$\Vert\mathbf{\Pi}_{\boldsymbol{V}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\Vert_{L^2(\Omega)}$		$\Vert\mathbf{\Pi}_{\boldsymbol{W}}\boldsymbol{u} - \boldsymbol{u}_h\Vert_{L^2(\Omega)}$	
		Mesh-1				Mesh-2			
k	Mesh	Error	Order	Error	Order	Error	Order	Error	Order
	h	9.81E-02	-	3.74E-03	-	4.20E-01	-	8.28E+12	-
0	$h/2$	9.50E-02	0.05	3.69E-02	0.02	2.14E-02	0.97	1.05E+12	2.98
	$h/4$	9.42E-02	0.01	3.68E-03	0.00	1.05E-02	1.02	1.41E+11	2.90
	$h/8$	9.41E-02	0.00	3.68E-03	0.00	5.22E-03	1.01	1.79E+10	2.98
	$h/16$	9.41E-02	0.00	3.68E-03	0.00	8.17E-01	-3.97	1.39E+12	-6.28
	h	2.26E-03	-	1.88E-03	-	2.04E-03	-	9.41E-04	-
1	$h/2$	7.24E-03	1.65	3.57E-04	2.40	5.90E-03	1.79	1.45E-04	2.70
	$h/4$	2.09E-03	1.79	5.51E-05	2.69	1.58E-03	1.92	2.00E-05	2.86
	$h/8$	5.60E-04	1.90	7.60E-06	2.86	4.08E-04	1.95	2.62E-06	2.93
	$h/16$	1.45E-04	1.95	9.93E-07	2.94	7.01E-06	2.00	3.35E-07	2.97
	h	1.24E-03	-	5.52E-05	-	1.23E-03	-	3.53E-05	-
2	$h/2$	1.57E-04	2.98	3.74E-06	3.88	1.57E-04	2.97	2.25E-06	3.97
	$h/4$	1.97E-05	2.99	2.43E-07	3.95	1.97E-05	2.99	1.42E-07	3.99
	$h/8$	2.46E-06	3.00	1.54E-08	3.97	2.47E-06	3.00	8.90E-09	3.99
	$h/16$	3.08E-07	3.00	9.73E-10	3.99	3.10E-07	3.00	5.58E-10	4.00
	h	5.26E-05	-	1.45E-06	-	5.33E-05	-	1.27E-06	-
3	$h/2$	3.51E-06	3.90	4.90E-08	4.89	3.54E-06	3.91	4.36E-08	4.86
	$h/4$	2.26E-07	3.96	1.59E-09	4.95	2.29E-07	3.95	1.43E-09	4.93
	$h/8$	1.42E-08	3.98	5.12E-11	4.96	1.45E-08	3.98	4.59E-11	4.96

5.2. **Locking experiments.** In this section, we consider an isotropic material in 2D with plane-strain condition:

$$\mathcal{A}\underline{\boldsymbol{\sigma}} = \frac{1+\nu}{E}\underline{\boldsymbol{\sigma}} - \frac{(1+\nu)\nu}{E}\mathrm{tr}(\underline{\boldsymbol{\sigma}})I_2, \tag{5.3}$$

where ν is the Poisson Ratio and E is the Young's Modulus. This example satisfies Assumption 2.1 with $P_D = \frac{1+\nu}{E}$ and $P_T = \frac{(1+\nu)}{E}(1-2\nu)$. By sending $\nu \to 0.5$, this material is nearly incompressible. We consider an example in [10, 35] by setting $\boldsymbol{f}$ and $\boldsymbol{g}$ to satisfy the exact solution:

$$u_1 = -x^2(x-1)^2 y(y-1)(2y-1), \tag{5.4}$$

$$u_2 = y^2(y-1)^2 x(x-1)(2x-1), \tag{5.5}$$

with $E = 3$. We conduct numerical experiments for this problem for $k = 1, 2, 3$ with $\tau = \mathcal{O}(\frac{1}{h})$. The history of convergence is displayed in Table 3 and the convergence sequence of the stress and the displacement is plotted in Figure 3. By increasing ν from 0.49 to 0.49999, we observe the same order of convergence which is optimal in both stress and displacement. In addition, our numerical results demonstrate that the convergence properties of our method do not depend on the type of meshes. Altogether, this observation exactly aligns with the error estimates provided in Theorem 2.2 and it justifies that our HDG method is free from volumetric-locking.

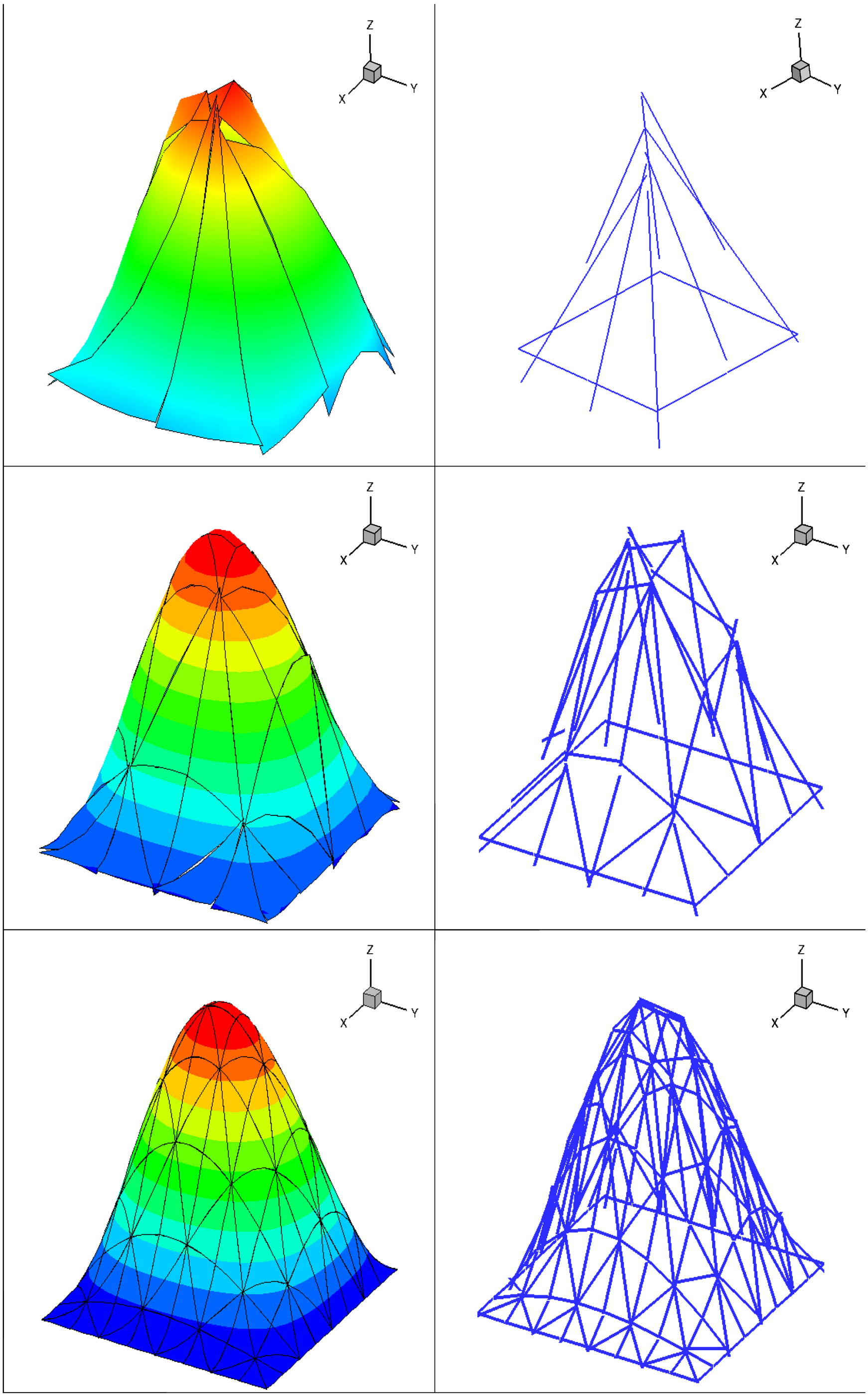

FIGURE 2. Convergence sequence of the displacement on Mesh-2 for $k = 1$. Left: u_h^1(quadratic), Right: $\widehat{u}_h^1$(linear)

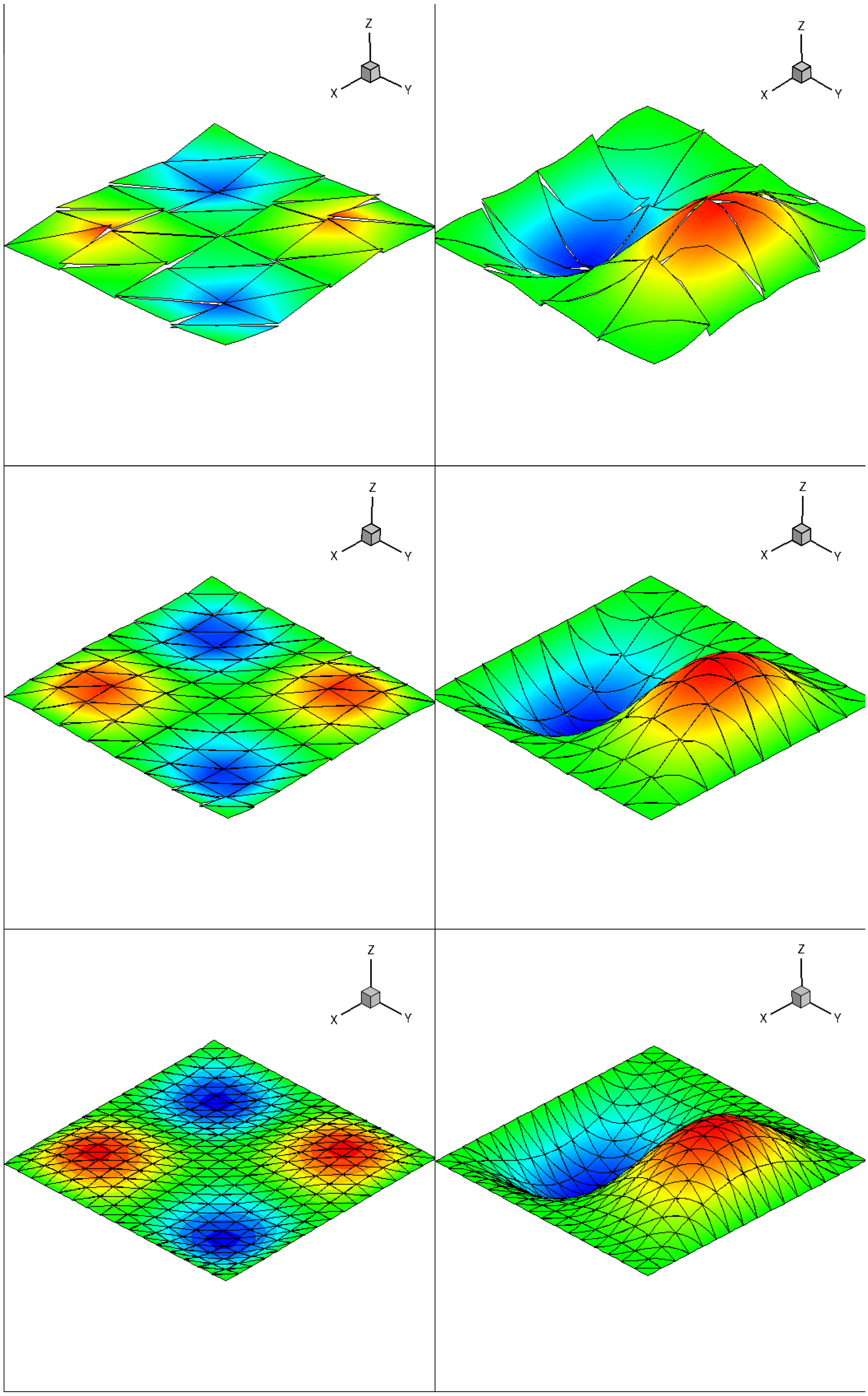

FIGURE 3. Convergence sequence of the stress and the displacement on Mesh-1 for $k = 1$ and $\nu = 0.49999$. Left: σ_h^{11}(linear), right u_h^1(quadratic).

TABLE 3. History of convergence for the exact solution (5.4) where $h = 0.354$.

		$\|\underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{L^2(\Omega)}$		$\|\boldsymbol{\Pi_W u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$		$\|\underline{\boldsymbol{\Pi_V}\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{L^2(\Omega)}$		$\|\boldsymbol{\Pi_W u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$	
				$\nu = 0.49$					
		Mesh-1				Mesh-2			
k	Mesh	Error	Order	Error	Order	Error	Order	Error	Order
	h	4.12E-03	-	1.14E-04	-	4.12E-03	-	9.15E-04	-
1	$h/2$	1.22E-03	1.75	2.53E-05	2.17	1.27E-03	1.70	1.47E-05	2.64
	$h/4$	3.32E-04	1.88	4.76E-06	2.41	3.40E-04	1.90	2.00E-06	2.87
	$h/8$	8.69E-05	1.93	8.17E-07	2.54	8.64E-05	1.98	2.58E-07	2.96
	$h/16$	2.22E-05	1.97	1.23E-07	2.73	2.17E-05	1.99	3.27E-08	2.98
	h	9.33E-04	-	2.00E-05	-	9.37E-04	-	1.24E-05	-
2	$h/2$	1.29E-04	2.85	1.65E-06	3.60	1.32E-04	2.83	9.42E-07	3.71
	$h/4$	1.65E-05	2.97	1.17E-07	3.82	1.64E-05	3.00	6.01E-08	3.97
	$h/8$	2.07E-06	3.00	7.76E-09	3.92	2.05E-06	3.00	3.77E-09	3.99
	$h/16$	2.58E-07	3.00	4.98E-10	3.96	2.56E-07	3.00	2.36E-10	4.00
	h	1.44E-04	-	1.65E-06	-	1.57E-04	-	1.53E-06	-
3	$h/2$	9.78E-06	3.88	6.18E-08	4.74	9.87E-06	3.99	5.11E-08	4.90
	$h/4$	6.27E-07	3.96	2.09E-09	4.89	6.19E-07	3.99	1.68E-09	4.93
	$h/8$	3.95E-08	3.99	6.77E-11	4.95	3.89E-08	3.99	5.43E-11	4.95
	$h/16$	2.49E-09	3.99	2.21E-12	4.94	2.44E-09	4.00	1.74E-12	4.97
				$\nu = 0.4999$					
		Mesh-1				Mesh-2			
k	Mesh	Error	Order	Error	Order	Error	Order	Error	Order
	h	4.12E-03	-	1.13E-04	-	4.13E-03	-	9.05E-04	-
1	$h/2$	1.22E-03	1.76	2.52E-05	2.17	1.26E-03	1.71	1.45E-05	2.64
	$h/4$	3.31E-04	1.88	4.72E-06	2.41	3.39E-04	1.90	1.98E-06	2.87
	$h/8$	8.66E-05	1.93	8.11E-07	2.54	8.61E-05	1.98	2.55E-07	2.96
	$h/16$	2.21E-05	1.97	1.22E-07	2.73	2.16E-05	1.99	3.23E-08	2.98
	h	9.32E-04	-	1.98E-05	-	9.34E-04	-	1.22E-05	-
2	$h/2$	1.29E-04	2.86	1.64E-06	3.60	1.32E-04	2.83	9.31E-07	3.72
	$h/4$	1.64E-05	2.97	1.16E-07	3.82	1.64E-05	3.00	5.94E-08	3.97
	$h/8$	2.06E-06	3.00	7.70E-09	3.92	2.04E-06	3.00	3.73E-09	3.99
	$h/16$	2.57E-07	3.00	4.95E-10	3.96	2.55E-07	3.00	2.33E-10	4.00
	h	1.44E-04	-	1.63E-06	-	1.57E-04	-	1.51E-06	-
3	$h/2$	9.75E-06	3.88	6.09E-08	4.74	9.83E-06	3.99	5.03E-08	4.90
	$h/4$	6.25E-07	3.96	2.06E-09	4.89	6.17E-07	3.99	1.66E-09	4.93
	$h/8$	3.94E-08	3.99	6.72E-11	4.95	3.87E-08	3.99	5.39E-11	4.94
	$h/16$	2.48E-09	3.99	2.20E-12	4.93	2.43E-09	3.99	1.73E-12	4.96
				$\nu = 0.49999$					
		Mesh-1				Mesh-2			
k	Mesh	Error	Order	Error	Order	Error	Order	Error	Order
	h	4.12E-03	-	1.13E-04	-	4.13E-03	-	9.05E-04	-
1	$h/2$	1.22E-03	1.76	2.52E-05	2.17	1.26E-03	1.71	1.45E-05	2.64
	$h/4$	3.31E-04	1.88	4.72E-06	2.41	3.39E-04	1.90	1.98E-06	2.87
	$h/8$	8.66E-05	1.93	8.11E-07	2.54	8.61E-05	1.98	2.55E-07	2.96
	$h/16$	2.21E-05	1.97	1.22E-07	2.73	2.16E-05	1.99	3.23E-08	2.98
	h	9.32E-04	-	1.98E-05	-	9.34E-04	-	1.22E-05	-
2	$h/2$	1.29E-04	2.86	1.64E-06	3.60	1.32E-04	2.83	9.31E-07	3.72
	$h/4$	1.64E-05	2.97	1.16E-07	3.82	1.64E-05	3.00	5.94E-08	3.97
	$h/8$	2.06E-06	3.00	7.70E-09	3.92	2.04E-06	3.00	3.73E-09	3.99
	$h/16$	2.57E-07	3.00	4.95E-10	3.96	2.55E-07	3.00	2.33E-10	4.00
	h	1.44E-04	-	1.63E-06	-	1.57E-04	-	1.50E-06	-
3	$h/2$	9.75E-06	3.88	6.09E-08	4.74	9.83E-06	3.99	5.03E-08	4.90
	$h/4$	6.25E-07	3.96	2.06E-09	4.89	6.17E-07	3.99	1.66E-09	4.93
	$h/8$	3.94E-08	3.99	6.72E-11	4.95	3.86E-08	3.99	5.39E-11	4.94
	$h/16$	2.48E-09	3.99	2.20E-12	4.93	2.44E-09	3.98	1.74E-12	4.95

ACKNOWLEDGEMENTS

The work of the first author was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11302014). As a convention the names of the authors are alphabetically ordered. All authors contributed equally in this article. Finally the authors would like to thank Guosheng Fu at the University of Minnesota for a fruitful discussion.

References

[1] S. Adams and B. Cockburn, *A mixed finite element method for elasticity in three dimensions*, J. Sci. Comput. **25** (2005), no. 3, 515–521. MR2221175

[2] D. N. Arnold, G. Awanou, and R. Winther, *Finite elements for symmetric tensors in three dimensions*, Math. Comp. **77** (2008), no. 263, 1229–1251. MR2398766

[3] D. N. Arnold, G. Awanou, and R. Winther, *Nonconforming tetrahedral mixed finite elements for elasticity*, Math. Models Methods Appl. Sci. **24** (2014), no. 4, 783–796. MR3163400

[4] D. N. Arnold, F. Brezzi, and J. Douglas Jr., *PEERS: a new mixed finite element for plane elasticity*, Japan J. Appl. Math. **1** (1984), no. 2, 347–367. MR840802

[5] D. N. Arnold, R. S. Falk, and R. Winther, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Math. Comp. **76** (2007), no. 260, 1699–1723. MR2336264

[6] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numer. Math. **92** (2002), no. 3, 401–419. MR1930384

[7] D. N. Arnold and R. Winther, *Nonconforming mixed elements for elasticity*, Math. Models Methods Appl. Sci. **13** (2003), no. 3, 295–307. MR1977627

[8] G. Awanou, *A rotated nonconforming rectangular mixed element for elasticity*, Calcolo **46** (2009), no. 1, 49–60. MR2495247

[9] C. Bacuta and J. H. Bramble, *Regularity estimates for solutions of the equations of linear elasticity in convex plane polygonal domains*, Z. Angew. Math. Phys. **54** (2003), no. 5, 874–878. MR2019187

[10] M. Bercovier and E. Livne, *A* 4 *CST quadrilateral element for incompressible materials and nearly incompressible materials*, Calcolo **16** (1979), no. 1, 5–19. MR555452

[11] D. Boffi, F. Brezzi, and M. Fortin, *Reduced symmetry elements in linear elasticity*, Commun. Pure Appl. Anal. **8** (2009), no. 1, 95–121. MR2449101

[12] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Series in Computational Mathematics, vol. 44, Springer, Heidelberg, 2013. MR3097958

[13] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991. MR1115205

[14] P. G. Ciarlet, *On Korn's inequality*, Chin. Ann. Math. Ser. B **31** (2010), no. 5, 607–618. MR2726058

[15] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. **77** (2008), no. 264, 1887–1916. MR2429868

[16] B. Cockburn, O. Dubois, J. Gopalakrishnan, and S. Tan, *Multigrid for an HDG method*, IMA J. Numer. Anal. **34** (2014), no. 4, 1386–1425. MR3269430

[17] B. Cockburn, J. Gopalakrishnan, and J. Guzmán, *A new elasticity element made for enforcing weak stress symmetry*, Math. Comp. **79** (2010), no. 271, 1331–1349. MR2629995

[18] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. **47** (2009), no. 2, 1319–1365. MR2485455

[19] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp. **79** (2010), no. 271, 1351–1367. MR2629996

[20] B. Cockburn, W. Qiu, and K. Shi, *Conditions for superconvergence of HDG methods for second-order elliptic problems*, Math. Comp. **81** (2012), no. 279, 1327–1353. MR2904581

[21] B. Cockburn, D. Schötzau, and J. Wang, *Discontinuous Galerkin methods for incompressible elastic materials*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 25-28, 3184–3204. MR2220915

[22] B. Cockburn and K. Shi, *Superconvergent HDG methods for linear elasticity with weakly symmetric stresses*, IMA J. Numer. Anal. **33** (2013), no. 3, 747–770. MR3081483

[23] G. Fu, B. Cockburn, and H. Stolarski, *Analysis of an HDG method for linear elasticity*, Internat. J. Numer. Methods Engrg. **102** (2015), no. 3-4, 551–575. MR3340089

[24] L. Gastaldi and R. H. Nochetto, *Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations* (English, with French summary), RAIRO Modél. Math. Anal. Numér. **23** (1989), no. 1, 103–128. MR1015921

[25] J. Gopalakrishnan and J. Guzmán, *A second elasticity element using the matrix bubble*, IMA J. Numer. Anal. **32** (2012), no. 1, 352–372. MR2875255

[26] J. Gopalakrishnan and J. Guzmán, *Symmetric nonconforming mixed finite elements for linear elasticity*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1504–1520. MR2831058

[27] J. Gopalakrishnan and W. Qiu, *An analysis of the practical DPG method*, Math. Comp. **83** (2014), no. 286, 537–552. MR3143683

[28] J. Guzmán, *A unified analysis of several mixed methods for elasticity with weak stress symmetry*, J. Sci. Comput. **44** (2010), no. 2, 156–169. MR2659794

[29] C. Lehrenfeld, *Hybrid Discontinuous Galerkin methods for solving incompressible flow problems*, Ph.D. Thesis (2010).

[30] H.-Y. Man, J. Hu, and Z.-C. Shi, *Lower order rectangular nonconforming mixed finite element for the three-dimensional elasticity problem*, Math. Models Methods Appl. Sci. **19** (2009), no. 1, 51–65. MR2484491

[31] I. Oikawa, *A hybridized discontinuous Galerkin method with reduced stabilization*, J. Sci. Comput. **65** (2015), no. 1, 327–340. MR3394448

[32] W. Qiu and L. Demkowicz, *Mixed hp-finite element method for linear elasticity with weakly imposed symmetry*, Comput. Methods Appl. Mech. Engrg. **198** (2009), no. 47-48, 3682–3701. MR2557491

[33] W. Qiu and L. Demkowicz, *Mixed hp-finite element method for linear elasticity with weakly imposed symmetry: stability analysis*, SIAM J. Numer. Anal. **49** (2011), no. 2, 619–641. MR2792388

[34] S.-C. Soon, *Hybridizable discontinuous Galerkin methods for solid mechanics*, Ph.D. Thesis, University of Minnesota, 2008.

[35] S.-C. Soon, B. Cockburn, and H. K. Stolarski, *A hybridizable discontinuous Galerkin method for linear elasticity*, Internat. J. Numer. Methods Engrg. **80** (2009), no. 8, 1058–1092. MR2589528

[36] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math. **53** (1988), no. 5, 513–538. MR954768

[37] S.-Y. Yi, *Nonconforming mixed finite element methods for linear elasticity using rectangular elements in two and three dimensions*, Calcolo **42** (2005), no. 2, 115–133. MR2158594

[38] S.-Y. Yi, *A new nonconforming mixed finite element method for linear elasticity*, Math. Models Methods Appl. Sci. **16** (2006), no. 7, 979–999. MR2250030

[39] N. C. Nguyen and J. Peraire, *Hybridizable discontinuous Galerkin methods for partial differential equations in continuum mechanics*, J. Comput. Phys. **231** (2012), no. 18, 5955–5988. MR2954837

Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Hong Kong

E-mail address: weifeqiu@cityu.edu.hk

Department of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

E-mail address: shenx179@umn.edu

Department of Mathematics & Statistics, Old Dominion University, Norfolk, Virginia 23529

E-mail address: kshi@odu.edu