

# Analysis of a Staggered Discontinuous Galerkin Method for Linear Elasticity

Jeonghun J. Lee<sup>1</sup> · Hyea Hyun Kim<sup>2</sup>

Received: 22 August 2014 / Revised: 20 April 2015 / Accepted: 22 April 2015 /  
Published online: 7 May 2015  
© European Union 2015

**Abstract** We develop a staggered discontinuous Galerkin method for linear elasticity problems and prove its a priori error estimates. In our variational formulation the symmetry of the stress tensor is imposed weakly via Lagrange multipliers but the resulting numerical stress tensor is strongly symmetric. Optimal a priori error estimates are obtained and the error estimates are robust in nearly incompressible materials. Numerical experiments illustrating our theoretical analysis are included.

**Keywords** Discontinuous Galerkin method · Linear elasticity · Error analysis · Weak symmetry

**Mathematics Subject Classification** 65N30 · 65N12

## 1 Introduction

Let  $\Omega$  be a bounded domain in  $\mathbb{R}^d$ ,  $d = 2, 3$ , with a Lipschitz boundary  $\partial\Omega$ . Let  $\mathbb{M}$ ,  $\mathbb{S}$ ,  $\mathbb{K}$  be the sets of all, symmetric, skew-symmetric  $d \times d$  matrices, and  $\mathbb{V}$  be the set of (column)  $\mathbb{R}^d$ -vectors. In (static) linear elasticity problems, we find mechanical quantities of a linearly elastic body occupying  $\Omega$  for a given body force and a given boundary condition. The problem

---

Jeonghun J. Lee was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) ERC Grant agreement 339643. Hyea Hyun Kim was supported by the Korea Science and Engineering Foundation (KOSEF) Grant funded by the Korean Government(MOST) (2011-0023354).

---

✉ Jeonghun J. Lee  
jeonghul@math.uio.no  
Hyea Hyun Kim  
hhkim@khu.ac.kr

<sup>1</sup> Department of Mathematics, University of Oslo, 0316 Oslo, Norway

<sup>2</sup> Department of Applied Mathematics, Kyung Hee University, Yongin, South Korea

can be written as a system of equations seeking two unknowns  $\sigma$  and  $u$ , the stress tensor and the displacement field, such that

$$A\sigma = \epsilon(u), \quad -\operatorname{div} \sigma = f \quad \text{in } \Omega. \quad (1)$$

Here  $\sigma$  and  $u$  are  $\mathbb{S}$ -valued and  $\mathbb{V}$ -valued functions on  $\Omega$ , respectively,  $A$  is the compliance tensor determined by material parameters of the elastic medium,  $\epsilon(u)$  is the symmetric part of the gradient of  $u$ ,  $f : \Omega \rightarrow \mathbb{V}$  is a given body force, and  $\operatorname{div}$  is the row-wise divergence. If the elastic body is clamped on the boundary of  $\Omega$ , then  $u = 0$  on  $\partial\Omega$  is imposed as a boundary condition.

For the compliance tensor  $A$ , we assume that  $A(x)$  for each  $x \in \Omega$  is a linear symmetric positive definite map from  $\mathbb{S}$  to  $\mathbb{S}$ , and there exist uniform upper and lower bounds of  $A(x)$  which are independent of  $x$ . For  $\tau : \Omega \rightarrow \mathbb{S}$ ,  $A\tau$  naturally gives an  $\mathbb{S}$ -valued function on  $\Omega$ . In a homogeneous isotropic elastic medium,  $A$  has the form

$$A\tau = \frac{1}{2\mu} \left( \tau - \frac{\lambda}{2\mu + d\lambda} \operatorname{tr}(\tau)\mathbf{I} \right), \quad (2)$$

in which  $\mu, \lambda$  are positive constants, called the Lamé parameters,  $\operatorname{tr}(\tau)$  is the trace of function  $\tau : \Omega \rightarrow \mathbb{S}$ , and  $\mathbf{I}$  is the  $d \times d$  identity matrix. We refer to [20] for more details on properties of compliance tensors.

A natural finite element approach for (1) is the mixed method [12]. However, this mixed method turns out to be highly nontrivial due to the symmetry constraint of stress tensor, which is originally from the law of conservation of angular momentum [20]. There are known mixed finite elements for the problem with symmetric stress tensor but those elements usually need sophisticated high order polynomials for shape functions of the stress tensor [2, 5, 7, 22, 23]. An alternative approach for the problem, say weak symmetry approach, is to impose the symmetry constraint weakly [36]. Various mixed finite elements have been developed based on it [1, 4, 6, 10, 14, 15, 17, 19, 25, 26, 28, 32, 33] but numerical stress tensor in these methods is only weakly symmetric. Mixed discontinuous Galerkin (DG) methods have also been considered but it seems that there are only a few of such results [15, 30].

In this paper, for the linear elasticity Eq. (1), we will develop a staggered discontinuous Galerkin (SDG) method which has been successfully developed and analyzed for the acoustic wave equation and the Stokes Eq. [13, 24]. In SDG methods, the flux condition across the inter-element boundary can be naturally obtained by the staggered continuity property of the finite element functions, whereas an artificial numerical flux needs to be introduced in standard DG methods and other mixed DG methods [15, 30].

In order to deal with the symmetry constraint of stress tensor, we establish our SDG method using a variational formulation with the weak symmetry approach [36]. However, the numerical stress tensor obtained in our SDG method becomes strongly symmetric. Polynomials of the same degree are used for all the finite element spaces and the optimal order of  $L^2$ -error estimates are proved for the given order of polynomials. In contrast, most other mixed and DG methods, which use the variational formulation with weakly symmetric stress, give only weakly symmetric numerical stress tensor and require more sophisticated shape functions to achieve the stability and optimal order of  $L^2$ -error estimates. We also prove that our SDG method does not suffer from the volumetric locking [8]. More precisely, for  $A$  of the form in (2), we show that the  $L^2$ -errors of all the unknowns in our method do not grow unboundedly as  $\lambda$  becomes arbitrarily large. We also point out that SDG methods are relevant for time-dependent problems [13], so it is reasonable to expect applications of our method to time-dependent elasticity problems.

We remark that another SDG method with symmetric stress tensor can also be developed by imposing symmetry constraint a priori to all shape functions of stress tensor. This is a reduced method in a sense because only two unknowns, stress tensor and displacement, are used and the total number of degrees of freedom (DOFs) in this method is smaller than that of our SDG method. However, there are prices to pay. In this reduced method, symmetric functions with local  $H(\text{div})$ -continuity are used as shape functions for stress tensor. For error analysis, these symmetric shape functions require construction of new interpolation operator having the symmetric shape functions as its range because a standard (locally)  $H(\text{div})$  interpolation operator does not have range of symmetric functions. The construction of new interpolation operator is nontrivial and makes error analysis more complicated. In contrast, no new interpolation operator is needed in our SDG method. We only use the interpolation operators of standard finite elements [11] and of the SDG method for mixed Poisson problem [13]. In addition to this drawback from new interpolation operator for error analysis, there is a difficulty in implementation. In order to implement the reduced method, we need local DOFs for the symmetric locally  $H(\text{div})$  shape functions but finding such local DOFs is by no means straightforward. These drawbacks seem to outweigh the benefit from less number of DOFs, so we do not pursue the reduced method further. In our case, finite element functions for stress tensor and the additional Lagrange multipliers for the weak symmetry are discontinuous across the element boundaries and thus in practical implementation those unknowns can be eliminated element-wise to obtain a reduced system on the displacement.

This work is organized as follows. In Sect. 2, a variational formulation with weakly symmetric stress is introduced. In Sect. 3, we introduce an SDG method for the problem, prove its stability and a priori error estimates. In particular, robustness of the error estimates in nearly incompressible materials is proved. In Sect. 4, we present numerical results for test problems on structured/unstructured meshes and in nearly incompressible materials. Finally, auxiliary results for the proof of stability are provided in Sect. 1.

## 2 Variational Form with Weak Symmetry

Suppose that there are two open subsets  $\Lambda_D, \Lambda_N$  of  $\partial\Omega$  such that  $\Lambda_D \cap \Lambda_N = \emptyset, \overline{\Lambda_D} \cup \overline{\Lambda_N} = \partial\Omega$ . Suppose that  $\Lambda_D \neq \emptyset$  and a boundary condition of (1) is given by

$$u = g_D \quad \text{on } \Lambda_D, \quad \sigma n = g_N \quad \text{on } \Lambda_N, \tag{3}$$

with  $g_D : \Lambda_D \rightarrow \mathbb{V}, g_N : \Lambda_N \rightarrow \mathbb{V}$ , in which  $n : \partial\Omega \rightarrow \mathbb{V}$  is the unit outward normal vector field on  $\partial\Omega$  and  $\sigma n$  stands for the matrix-vector multiplication. The problem (1) with this boundary condition is well-posed under suitable regularity assumptions on  $g_D$  and  $g_N$ .

Here we introduce notation and some definitions. For a bounded open subset  $G \subset \mathbb{R}^d$  of Lipschitz boundary, we use  $L^2(G)$  to denote the Hilbert space of real-valued functions on  $G$  whose inner product and the corresponding norm are

$$(p, q)_G := \int_G pq \, dx, \quad \|p\|_{0,G} := (p, p)_G^{\frac{1}{2}}, \quad p, q \in L^2(G).$$

When  $G = \Omega$  or  $G$  is obvious in context, we will use  $(p, q)$  and  $\|p\|_0$  for simplicity. For  $r > 0, H^r(G)$  and  $\|\cdot\|_{r,G}$  stand for the standard Sobolev space based on the  $L^2(G)$ -norm with differentiability of order  $r$  [16] and the corresponding norm. For a nonnegative integer  $k, \mathcal{P}_k(G)$  is the space of polynomials on  $G$  of order  $\leq k$ . For a finite dimensional inner product space  $\mathbb{X}$ , all these function spaces are naturally extended to  $\mathbb{X}$ -valued functions

with counterparts  $L^2(G; \mathbb{X})$ ,  $H^r(G; \mathbb{X})$ , and  $\mathcal{P}_k(G; \mathbb{X})$ . We will use same notation for inner products and norms in these extended function spaces, and the meaning will be clear in context. Let

$$H(\operatorname{div}, G) := \left\{ v \in L^2(G; \mathbb{R}^d) : \operatorname{div} v \in L^2(G) \right\},$$

in which  $\operatorname{div} v$  is the divergence of  $v$  in the sense of distributions. Recall that  $\operatorname{div}$  is regarded as row-wise divergence for  $\mathbb{M}$ -valued functions. Now we define

$$H(\operatorname{div}, G; \mathbb{M}) := \left\{ \boldsymbol{\tau} \in L^2(G; \mathbb{M}) : \operatorname{div} \boldsymbol{\tau} \in L^2(G; \mathbb{V}) \right\}.$$

For  $v \in H^1(G; \mathbb{V})$ , the grad operator stands for the row-wise gradient and thus  $\operatorname{grad} v \in L^2(G; \mathbb{M})$ . For an  $\mathbb{M}$ -valued function  $\boldsymbol{\tau}$ , the symmetric and skew-symmetric parts of  $\boldsymbol{\tau}$  are

$$\operatorname{sym} \boldsymbol{\tau} = \frac{\boldsymbol{\tau} + \boldsymbol{\tau}^T}{2}, \quad \operatorname{skw} \boldsymbol{\tau} = \frac{\boldsymbol{\tau} - \boldsymbol{\tau}^T}{2},$$

where  $\boldsymbol{\tau}^T$  is the transpose of  $\boldsymbol{\tau}$ . Note that  $\epsilon(u) = \operatorname{sym} \operatorname{grad} u$  by these definitions. Finally, for a bounded subset  $G \subset \mathbb{R}^d$  of codimension 1,  $L^2(G)$  is the Hilbert space of square-integrable functions on  $G$  with inner products

$$\langle p, q \rangle_G := \int_G pq \, ds, \quad p, q \in L^2(G),$$

and other spaces  $H^r(G)$ ,  $L^2(G; \mathbb{X})$ ,  $H^r(G; \mathbb{X})$ ,  $\mathcal{P}_k(G)$ ,  $\mathcal{P}_k(G, \mathbb{X})$  and norms  $\|\cdot\|_{0,G}$ ,  $\|\cdot\|_{r,G}$  are similarly defined.

In order to derive a variational form of (1) with weakly imposed symmetry of the stress, we first extend the compliance tensor  $A(x)$  for  $x \in \Omega$  to be a map from  $\mathbb{M}$  into  $\mathbb{M}$  by extending  $A(x)$  to be a positive multiple of the identity map on  $\mathbb{K}$ . For simplicity we again denote this extended compliance tensor by  $A$  because we will not use the original compliance tensor. Note that  $A\boldsymbol{\sigma} = \epsilon(u)$  still holds because  $\boldsymbol{\sigma}$  is symmetric. We then introduce a new unknown  $\boldsymbol{\gamma} \in L^2(\Omega; \mathbb{K})$  instead of  $\operatorname{skw} \operatorname{grad} u$ , and rewrite (1) as

$$A\boldsymbol{\sigma} = \operatorname{grad} u - \boldsymbol{\gamma}, \quad -\operatorname{div} \boldsymbol{\sigma} = f, \tag{4}$$

with symmetry constraint  $(\boldsymbol{\sigma}, \boldsymbol{\eta}) = 0, \forall \boldsymbol{\eta} \in L^2(\Omega; \mathbb{K})$ . If the triple  $(\boldsymbol{\sigma}, u, \boldsymbol{\gamma})$  satisfies (4) with the symmetry constraint, then one can check that the pair  $(\boldsymbol{\sigma}, u)$  satisfies (1). To write this system with (3) as a variational form, set

$$\begin{aligned} \Sigma &= L^2(\Omega; \mathbb{M}), & \Gamma &= L^2(\Omega; \mathbb{K}), \\ U_{g_D} &= \{v \in H^1(\Omega; \mathbb{V}) : v|_{A_D} = g_D\}, \\ U_0 &= \{v \in H^1(\Omega; \mathbb{V}) : v|_{A_D} = 0\}. \end{aligned}$$

The first equation in (4) is simply  $(A\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\operatorname{grad} u - \boldsymbol{\gamma}, \boldsymbol{\tau})$  for all  $\boldsymbol{\tau} \in \Sigma$ . For  $v \in U_0$ , by the integration by parts, we have

$$(f, v) = -(\operatorname{div} \boldsymbol{\sigma}, v) = -\langle \boldsymbol{\sigma}n, v \rangle_{\partial\Omega} + (\boldsymbol{\sigma}, \operatorname{grad} v) = -\langle g_N, v \rangle_{A_N} + (\boldsymbol{\sigma}, \operatorname{grad} v),$$

where  $n$  is the outward unit normal vector field on  $\partial\Omega$ . Combining it with the symmetry constraint  $(\boldsymbol{\sigma}, \boldsymbol{\eta}) = 0$  for all  $\boldsymbol{\eta} \in \Gamma$ , we have

$$(\boldsymbol{\eta} - \operatorname{grad} v, \boldsymbol{\sigma}) = -(f, v) - \langle g_N, v \rangle_{A_N}, \quad \forall v \in U_0.$$

Rewriting  $u$  by  $u = u_0 + u_g$  with a fixed  $u_g \in U_{g_D}$ , a variational form is to find  $(\sigma, u_0, \gamma) \in \Sigma \times U_0 \times \Gamma$  such that

$$(A\sigma, \tau) + (\gamma - \text{grad } u_0, \tau) = (\text{grad } u_g, \tau), \quad \forall \tau \in \Sigma, \tag{5}$$

$$(\eta - \text{grad } v, \sigma) = -(f, v) - \langle g_N, v \rangle_{\Lambda_N}, \quad \forall (v, \eta) \in U_0 \times \Gamma. \tag{6}$$

For simplicity of presentation we will assume that  $g_D = g_N = 0$  with  $u_g = 0$  in the rest of the paper because it is straightforward to generalize our method, which will be described in the next section, to more general boundary conditions.

### 3 A Staggered DG Method, Stability, and Error Analysis

In this section we develop a staggered DG method for (5)–(6), prove its stability, and show a priori error estimates.

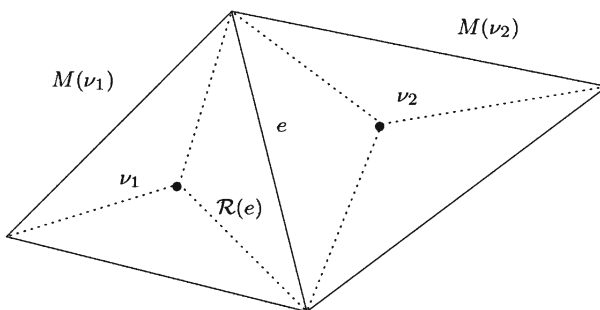
#### 3.1 A Staggered DG Method and Stability

Let  $\mathcal{M}_h$  be a shape-regular triangulation of  $\Omega$  with parameter  $h$ , the largest diameter of simplices in  $\mathcal{M}_h$ , and assume that  $\mathcal{M}_h$  is conforming to boundary subdomains  $\Lambda_D, \Lambda_N$ . Throughout this work,  $c$  will denote a positive constant independent of  $h$  but it can be different in each formula in order to avoid proliferation of constants in inequalities. Let  $\mathcal{T}_h$  be the simplicial mesh obtained from  $\mathcal{M}_h$  by refining  $\mathcal{M}_h$  with barycentric subdivision, and  $\mathcal{F}_h$  be the set of subsimplices of  $\mathcal{T}_h$  with codimension one. The sets of subsimplices  $\mathcal{F}_u, \mathcal{F}_\sigma$ , and  $\mathcal{F}_h^\circ$  are defined by

$$\mathcal{F}_u = \{e \in \mathcal{F}_h : e \subset \partial M, M \in \mathcal{M}_h\}, \quad \mathcal{F}_\sigma = \mathcal{F}_h \setminus \mathcal{F}_u, \quad \mathcal{F}_h^\circ = \{e \in \mathcal{F}_h : e \not\subset \partial\Omega\}.$$

For an illustration of a staggered mesh and its subsimplices we refer to Fig. 1. Suppose that  $\tau \in L^2(\Omega; \mathbb{M})$  and  $v \in L^2(\Omega; \mathbb{V})$ , whose restrictions on each  $T \in \mathcal{T}_h$  are in  $H^r(T; \mathbb{M}), H^r(T; \mathbb{V})$  with  $r > 1/2$ . We now define the jumps  $[[\tau n]]$  and  $[[v]]$  on each  $e \in \mathcal{F}_h$ . For  $e \in \mathcal{F}_h$  with  $e \subset \partial\Omega$ , we define

$$[[\tau n]]|_e := \tau|_e n, \quad [[v]]|_e := v|_e$$



**Fig. 1** An illustration of a staggered mesh in two dimensions:  $M(v_i)$  are triangles in  $\mathcal{M}_h$ ,  $v_i$  is the barycenter of  $M(v_i)$ , solid lines are edges of  $\mathcal{F}_u$ , dotted lines are edges of  $\mathcal{F}_\sigma$ , and  $\mathcal{R}(e)$  is the union of triangles sharing an edge  $e$  in  $\mathcal{F}_u$

where  $n$  is the unit outward normal vector on  $\partial\Omega$ , and  $\tau|_e, v|_e$  are the traces of  $\tau$  and  $v$  on  $e$ , respectively. To define the jumps on given  $e \in \mathcal{F}_h^\circ$ , let  $T^+, T^- \in \mathcal{T}_h$  be the unique two simplices such that  $e \subset \partial T^+ \cap \partial T^-$ , and we denote the unit normal vector on  $e$  pointing outward from  $T^+(T^-)$  by  $n^+(n^-)$ , respectively). We also denote by  $\tau^+$  the trace of  $\tau$  on  $e$  given by  $\tau|_{T^+}$ . The notations  $\tau^-, v^+, v^-$  are similarly defined. Then the jumps are defined by

$$[[\tau n]]|_e := \tau^+ n^+ + \tau^- n^-, \quad [[v]]|_e := v^+ - v^-. \tag{7}$$

Since the definition of  $[[v]]|_e$  depends on the choice of  $T^+$  and  $T^-$ , it will need an additional care in error analysis.

We are ready to define finite element spaces for the problem. The space  $\Sigma_h$  is the set of functions

$$\Sigma_h = \{ \tau \in \Sigma : \tau|_T \in \mathcal{P}_k(T; \mathbb{M}) \ \forall T \in \mathcal{T}_h \text{ and } [[\tau n]]|_e = 0 \ \forall e \in \mathcal{F}_\sigma \}, \quad k \geq 1.$$

The vanishing jump condition of  $\Sigma_h$  implies that the normal components of each row of  $\tau \in \Sigma_h$  are single-valued on every  $e \in \mathcal{F}_\sigma$ . The degrees of freedom of  $\Sigma_h$  are

$$\int_e \tau n_e \cdot p \, ds, \quad \forall p \in \mathcal{P}_k(e; \mathbb{V}), \quad \forall e \in \mathcal{F}_\sigma, \tag{8}$$

$$\int_T \tau : \phi \, dx, \quad \forall \phi \in \mathcal{P}_{k-1}(T; \mathbb{M}), \quad \forall T \in \mathcal{T}_h, \tag{9}$$

for  $\tau \in \Sigma_h$ , where the dot and colon are inner products on  $\mathbb{V}$  and  $\mathbb{M}$ , and  $n_e$  is a (chosen) unit normal vector on  $e$ . The well-definedness of  $\Sigma_h$  with respect to these DOFs is a consequence of [13, Lemma 2.3]. Let

$$\tilde{U}_h = \{ v \in L^2(\Omega; \mathbb{V}) : v|_T \in \mathcal{P}_k(T; \mathbb{V}) \ \forall T \in \mathcal{T}_h \text{ and } [[v]]|_e = 0 \ \forall e \in \mathcal{F}_u \},$$

and the jump condition of  $\tilde{U}_h$  implies that  $v \in \tilde{U}_h$  is continuous on every  $e \in \mathcal{F}_u$ . The degrees of freedom of  $\tilde{U}_h$  are

$$\int_e v \cdot p \, ds, \quad \forall p \in \mathcal{P}_k(e; \mathbb{V}), \quad \forall e \in \mathcal{F}_u, \tag{10}$$

$$\int_T v \cdot \phi \, dx, \quad \forall \phi \in \mathcal{P}_{k-1}(T; \mathbb{V}), \quad \forall T \in \mathcal{T}_h, \tag{11}$$

for  $v \in \tilde{U}_h$ , and we refer to [13, Lemma2.2] for its well-definedness. We define  $U_h$  as the subspace of  $\tilde{U}_h$  such that all the DOFs associated to  $e \in \mathcal{F}_u$  with  $e \subset \Lambda_D$  vanish.

By the construction,  $\tau$  in  $\Sigma_h$  is locally  $H(\text{div})$ -conforming in  $M(v_i)$  and  $v$  in  $U_h$  is locally  $H^1$ -conforming in  $\mathcal{R}(e)$  for  $e \in \mathcal{F}_\sigma$ , see Fig. 1. We observe that at least one of  $[[\tau n]]$  and  $[[v]]$  vanishes on  $e$  for any  $e \in \mathcal{F}_h^\circ$ . In other words, continuity of functions in the pair  $(\Sigma_h, U_h)$  is staggered on subsimplices in  $\mathcal{F}_h^\circ$ . This property will give us the inter-element flux condition in our Galerkin formulation without the need of carefully designed flux condition in other DG methods [3, 15, 30] while allowing partial discontinuity for the finite element spaces. The space  $\Gamma_h$  is defined by

$$\Gamma_h = \{ \eta \in \Gamma : \eta|_T \in \mathcal{P}_k(T; \mathbb{K}), \ \forall T \in \mathcal{T}_h \}.$$

We now define norms for  $\Sigma_h$  and  $U_h \times \Gamma_h$  by

$$\|\boldsymbol{\tau}\|_{0,h}^2 = \|\boldsymbol{\tau}\|_0^2 + \sum_{e \in \mathcal{F}_\sigma} h_e \|\boldsymbol{\tau}n\|_{0,e}^2, \tag{12}$$

$$\|(v, \boldsymbol{\eta})\|_h^2 = \|\boldsymbol{\eta} - \text{grad}_h v\|_0^2 + \sum_{e \in \mathcal{F}_\sigma} h_e^{-1} \|\llbracket v \rrbracket\|_{0,e}^2, \tag{13}$$

where  $\text{grad}_h$  is the element-wise grad operator with respect to  $\mathcal{T}_h$ , and  $h_e$  is the diameter of simplex  $e$ . To see that  $\|(\cdot, \cdot)\|_h$  is a norm, suppose that  $\|(v, \boldsymbol{\eta})\|_h = 0$  for  $v \in U_h, \boldsymbol{\eta} \in \Gamma_h$ . This is equivalent to  $\boldsymbol{\eta} - \text{grad}_h v = 0$  and  $\llbracket v \rrbracket|_e = 0$  for all  $e \in \mathcal{F}_\sigma$ . Considering the symmetric and skew-symmetric parts of  $\boldsymbol{\eta} - \text{grad}_h v$ , we have  $\text{sym grad}_h v = 0$  and  $\boldsymbol{\eta} - \text{skw grad}_h v = 0$ . Then  $v$  is an element-wise rigid body motion because  $\text{sym grad}_h v = 0$ . Furthermore, since  $\llbracket v \rrbracket|_e = 0$  for all  $e \in \mathcal{F}_\sigma$  and  $v$  is continuous on all  $e \in \mathcal{F}_u, v$  is a rigid body motion on  $\Omega$ . This implies that  $v = 0$  because  $v = 0$  on  $\Lambda_D \neq \emptyset$ . Then  $\boldsymbol{\eta} = 0$  as well because  $\boldsymbol{\eta} - \text{skw grad}_h v = 0$ . The other conditions of  $\|(\cdot, \cdot)\|_h$  to be a norm, are easy to check. For pure traction boundary conditions ( $\Lambda_D = \emptyset$ ), a modified argument is needed and we will discuss it later.

Recall that  $\llbracket v \rrbracket|_e$  in (7) and  $\boldsymbol{\tau}n|_e$  on  $e \in \mathcal{F}_\sigma$  depend on the choice of  $T^+, T^-$ , and the direction of unit normal vector  $n$  on  $e$ , respectively. To avoid this ambiguity in our analysis, whenever we deal with a term of the form  $\langle \llbracket v \rrbracket, \boldsymbol{\tau}n \rangle_e$  on  $e \in \mathcal{F}_\sigma$ , we assume that  $n$  is chosen to satisfy:

$$n = n^+ \text{ for the simplices } T^+ \text{ and } T^- \text{ such that } \llbracket v \rrbracket = v^+ - v^-.$$

Adopting this convention, let us define bilinear forms

$$\begin{aligned} a(\boldsymbol{\phi}, \boldsymbol{\tau}) &= (A\boldsymbol{\phi}, \boldsymbol{\tau}), \\ b(\boldsymbol{\tau}; v, \boldsymbol{\eta}) &= (\boldsymbol{\eta} - \text{grad}_h v, \boldsymbol{\tau}) + \sum_{e \in \mathcal{F}_\sigma} \langle \llbracket v \rrbracket, \boldsymbol{\tau}n \rangle_e, \\ b^*(v, \boldsymbol{\eta}; \boldsymbol{\tau}) &= (v, \text{div}_h \boldsymbol{\tau}) - \sum_{e \in \mathcal{F}_u} \langle v, \llbracket \boldsymbol{\tau}n \rrbracket \rangle_e + (\boldsymbol{\eta}, \boldsymbol{\tau}), \end{aligned} \tag{14}$$

for  $\boldsymbol{\phi}, \boldsymbol{\tau} \in \Sigma_h, v \in U_h, \boldsymbol{\eta} \in \Gamma_h$ , in which  $\text{div}_h$  is the element-wise div operator with respect to  $\mathcal{T}_h$ . In fact,  $b(\boldsymbol{\tau}; v, \boldsymbol{\eta})$  and  $b^*(v, \boldsymbol{\eta}; \boldsymbol{\tau})$  are well-defined for  $\boldsymbol{\tau} \in L^2(\Omega; \mathbb{M}), v \in L^2(\Omega; \mathbb{V}), \boldsymbol{\eta} \in L^2(\Omega; \mathbb{K})$  when  $\boldsymbol{\tau}$  and  $v$  are regular enough that  $\text{grad}_h v, \text{div}_h \boldsymbol{\tau}$  are meaningful and the integrations on  $e \in \mathcal{F}_h$  in the definitions are well-defined.

Notice that

$$\sum_{T \in \mathcal{T}_h} \langle v, \boldsymbol{\tau}n_T \rangle_{\partial T} = \sum_{e \in \mathcal{F}_\sigma} \langle \llbracket v \rrbracket, \boldsymbol{\tau}n \rangle_e + \sum_{e \in \mathcal{F}_u} \langle v, \llbracket \boldsymbol{\tau}n \rrbracket \rangle_e, \quad \forall v \in U_h, \forall \boldsymbol{\tau} \in \Sigma_h,$$

where  $n_T$  is the outward unit normal vector field on  $\partial T$ . This fact and the integration by parts give

$$b^*(v, \boldsymbol{\eta}; \boldsymbol{\tau}) = b(\boldsymbol{\tau}; v, \boldsymbol{\eta}), \quad \forall (\boldsymbol{\tau}, v, \boldsymbol{\eta}) \in \Sigma_h \times U_h \times \Gamma_h. \tag{15}$$

We now establish a discrete counterpart of the problem (5)–(6) in our numerical scheme: Find  $(\boldsymbol{\sigma}_h, u_h, \boldsymbol{\gamma}_h) \in \Sigma_h \times U_h \times \Gamma_h$  such that

$$a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + b^*(u_h, \boldsymbol{\gamma}_h; \boldsymbol{\tau}) = 0, \quad \forall \boldsymbol{\tau} \in \Sigma_h, \tag{16}$$

$$-b(\boldsymbol{\sigma}_h; v, \boldsymbol{\eta}) = (f, v), \quad \forall v \in U_h, \forall \boldsymbol{\eta} \in \Gamma_h. \tag{17}$$

We emphasize that in the above the flux condition on any  $e$  in  $\mathcal{F}_h$  is naturally obtained by the staggered continuity property of our finite element function spaces in contrast to other standard DG methods. For the consistency of this discretization, we show that the solution of (5)–(6) satisfies (16)–(17) under a suitable regularity assumption.

**Lemma 1** *For given  $f \in L^2(\Omega; \mathbb{V})$ , suppose that  $(\sigma, u, \gamma)$  is a solution of (5)–(6) with*

$$\sigma \in H^r(\Omega; \mathbb{S}), \quad u \in H^r(\Omega; \mathbb{V}), \quad r > 1/2. \tag{18}$$

*Then  $(\sigma, u, \gamma)$  satisfies (16)–(17).*

*Proof* Under the regularity assumptions on  $\sigma$  and  $u$ ,  $\llbracket \sigma n \rrbracket|_e$  and  $\llbracket u \rrbracket|_e$  are well-defined and vanish on every  $e \in \mathcal{F}_h^\circ$ . Using  $A\sigma = \text{grad } u - \gamma = \text{grad}_h u - \gamma$  and the integration by parts with boundary condition  $g_N = g_D = 0$ , we have

$$(A\sigma, \tau) = (\text{grad}_h u - \gamma, \tau) = -(u, \text{div}_h \tau) + \sum_{e \in \mathcal{F}_u} \langle u, \llbracket \tau n \rrbracket \rangle_e - (\gamma, \tau), \quad \forall \tau \in \Sigma_h,$$

so  $(\sigma, u, \gamma)$  satisfies (16). Similarly, if we use the fact  $(\sigma, \eta) = 0$  for all  $\eta \in \Gamma_h$ , the equality  $-\text{div } \sigma = f$ , and the integration by parts, then we have

$$(f, v) = -(\text{div } \sigma, v) - (\sigma, \eta) = (\sigma, \text{grad}_h v) - \sum_{e \in \mathcal{F}_\sigma} \langle \sigma n, \llbracket v \rrbracket \rangle_e - (\sigma, \eta),$$

and it implies that  $(\sigma, u, \gamma)$  satisfies (17). □

The stability of discretization, i.e., well-posedness of the problem (16)–(17), is obtained by the Babuška–Brezzi stability theory [12]. The following inf-sup condition, will be proved using the mesh-dependent norm idea in [34], is crucial in the stability theory. An auxiliary result Lemma 9, necessary in the inf-sup condition proof, is provided in the appendix.

**Lemma 2** *There exists  $\alpha > 0$  independent of  $h$  such that, for any  $(0, 0) \neq (v, \eta) \in U_h \times \Gamma_h$ , one can find  $0 \neq \tau \in \Sigma_h$  satisfying*

$$\frac{b(\tau; v, \eta)}{\|(v, \eta)\|_h \|\tau\|_{0,h}} \geq \alpha.$$

*Proof* For given  $v \in U_h$  and  $\eta \in \Gamma_h$ , suppose that  $b(\tau; v, \eta) = 0$  for all  $\tau \in \Sigma_h$ . For each  $M \in \mathcal{M}_h$ , we set

$$\mathcal{F}_\sigma^M = \{e \in \mathcal{F}_\sigma : e \subset M\}, \tag{19}$$

$$\Sigma_M = \{\tau \in \Sigma_h : \tau|_T = 0 \text{ for all } T \in \mathcal{T}_h \text{ such that } T \not\subset M\}, \tag{20}$$

$$\Sigma_{M,0} = \{\tau \in \Sigma_M : \tau n = 0 \text{ on } \partial M\}, \tag{21}$$

where  $n$  is the outward unit normal vector field on  $\partial M$ . For  $\tau \in \Sigma_M$ , the assumption  $b(\tau; v, \eta) = 0$  and the integration by parts give

$$\begin{aligned} 0 &= b(\tau; v, \eta) \\ &= (\eta - \text{grad}_h v, \tau)_M + \sum_{e \in \mathcal{F}_\sigma^M} \langle \llbracket v \rrbracket, \tau n \rangle_e \\ &= (\text{div}_h \tau, v)_M + (\tau, \eta)_M - \sum_{e \subset \partial M} \langle v, \llbracket \tau n \rrbracket \rangle_e. \end{aligned}$$



The condition on  $\tau n$  in the definition of  $\Sigma_{M,0}$  implies that  $[\tau n] = 0$  on  $\partial M$  for  $\tau \in \Sigma_{M,0}$ , so

$$(\operatorname{div}_h \tau, v)_M + (\tau, \eta)_M = 0, \quad \forall \tau \in \Sigma_{M,0}.$$

Let  $\eta_0$  be the component of  $\eta|_M$  such that each entry of  $\eta_0$  in  $\mathbb{K}$  is mean-value zero on  $M$ . By Lemma 9 in the appendix, one can find  $\tau_0 \in \Sigma_{M,0}$  such that  $\operatorname{div}_h \tau_0 = 0$  and  $\operatorname{skw} \tau_0|_M = \eta_0$ . Noting that  $(\tau_0, \eta)_M = (\operatorname{skw} \tau_0, \eta)_M = (\eta_0, \eta)_M = (\eta_0, \eta_0)_M$ , the above equation implies  $\eta_0 = 0$ , and therefore  $\eta|_M \in \mathcal{P}_0(M; \mathbb{K})$ . Recall that general forms of rigid body motions in  $\mathbb{R}^2$  and  $\mathbb{R}^3$  are

$$\begin{pmatrix} 0 & -s \\ s & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \begin{pmatrix} 0 & s_3 & -s_2 \\ -s_3 & 0 & s_1 \\ s_2 & -s_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad (22)$$

with constants  $s, s_i, a_i \in \mathbb{R}, i = 1, 2, 3$ . Since  $\eta|_M \in \mathcal{P}_0(M; \mathbb{K})$ , there is a rigid body motion  $r_M$  on  $M$  such that  $\eta|_M = \operatorname{grad} r_M$ . Then

$$b(\tau; v, \eta) = (\operatorname{grad}_h(r_M - v), \tau)_M + \sum_{e \in \mathcal{F}_\sigma^M} \langle [v], \tau n \rangle_e = 0.$$

Due to the polynomial degree of  $v$ , we have  $\operatorname{grad}_h(r_M - v)|_T \in \mathcal{P}_{k-1}(T; \mathbb{M})$  for each subsimplex  $T \subset M$  and  $[v]|_e \in \mathcal{P}_k(e; \mathbb{V})$  for  $e \in \mathcal{F}_\sigma^M$ . By (8) and (9), there is  $\tau \in \Sigma_M$  such that

$$\begin{aligned} \langle \tau n, [v] \rangle_e &= \|[v]\|_{0,e}^2, & \forall e \in \mathcal{F}_\sigma^M, \\ \int_T \tau : \operatorname{grad}_h(r_M - v) \, dx &= \int_T |\operatorname{grad}_h(r_M - v)|^2 \, dx, & \forall T \in \mathcal{T}_h \text{ with } T \subset M. \end{aligned}$$

Hence  $[v]|_e = 0$  for  $e \in \mathcal{F}_\sigma^M$  and  $\operatorname{grad}_h(r_M - v) = \eta - \operatorname{grad}_h v = 0$  on  $M$ .

If we do the same procedure for all  $M \in \mathcal{M}_h$ , then we can conclude that  $b(\tau; v, \eta) = 0$  for all  $\tau \in \Sigma_h$  implies  $\|(v, \eta)\|_h = 0$  by the definition of  $\|\cdot\|_h$  and the boundary condition  $v = 0$  on  $\Lambda_D$ . By equivalence of norms on finite dimensional spaces and a standard scaling argument, one can conclude that there is  $\alpha > 0$  independent of mesh sizes such that

$$\|\tau\|_{0,h} \leq \|(v, \eta)\|_h, \quad b(\tau; v, \eta) \geq \alpha \|(v, \eta)\|_h^2.$$

The assertion easily follows from this. □

**Theorem 1** *The problem (16)–(17) is well-posed.*

*Proof* Since  $A(x)$  is positive definite on  $\mathbb{M}$  uniformly in  $x \in \Omega$ , the bilinear form  $a(\cdot, \cdot)$  is coercive on  $L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{M})$ . By a standard scaling argument,  $\|\tau\|_0$  and  $\|\tau\|_{0,h}$  for  $\tau \in \Sigma_h$  are equivalent norms, so  $a(\cdot, \cdot)$  is also coercive on  $\Sigma_h \times \Sigma_h$ . By the Cauchy–Schwarz inequality,  $a(\cdot, \cdot)$  is a bounded bilinear form on  $\Sigma_h \times \Sigma_h$ . Again by the Cauchy–Schwarz inequality,

$$b(\tau; v, \eta) \leq \|\tau\|_{0,h} \|(v, \eta)\|_h, \quad \forall (\tau, v, \eta) \in \Sigma_h \times U_h \times \Gamma_h, \quad (23)$$

so  $b(\cdot; \cdot, \cdot)$  is a bounded bilinear form on  $\Sigma_h \times (U_h \times \Gamma_h)$ . Then the well-posedness of (16)–(17) follows by the inf-sup condition in Lemma 2, the equality (15), and the Babuška–Brezzi stability theory [12]. □

### 3.2 Error Analysis

Let  $(\sigma, u, \gamma)$  and  $(\sigma_h, u_h, \gamma_h)$  be the solutions of (5)–(6) and (16)–(17), respectively. Under the regularity assumption of  $\sigma$  and  $u$  in (18), the Babuška–Brezzi stability theory immediately gives

$$\begin{aligned} & \|\sigma - \sigma_h\|_{0,h} + \|(u - u_h, \gamma - \gamma_h)\|_h \\ & \leq c \inf_{(\tau, v, \eta) \in \Sigma_h \times U_h \times \Gamma_h} (\|\sigma - \tau\|_{0,h} + \|(u - v, \gamma - \eta)\|_h). \end{aligned} \tag{24}$$

However, orders of all the errors in this estimate are  $O(h^k)$  for sufficiently regular solutions, so this estimate is not optimal because the optimal  $L^2$ -error of  $\sigma$  is  $O(h^{k+1})$  with respect to the full approximability of finite element spaces. Therefore, we present an error analysis which gives optimal error bounds.

We first introduce some interpolation operators. For  $\tau \in H^r(\Omega; \mathbb{M})$ ,  $r > 1/2$ , let  $J$  be the interpolation operator defined by (8)–(9), i.e.,

$$\begin{aligned} \int_e (J\tau)n \cdot p \, ds &= \int_e \tau n \cdot p \, ds, \quad \forall p \in \mathcal{P}_k(e; \mathbb{V}), \quad \forall e \in \mathcal{F}_\sigma, \\ \int_T J\tau : \phi \, dx &= \int_T \tau : \phi \, dx, \quad \forall \phi \in \mathcal{P}_{k-1}(T; \mathbb{M}), \quad \forall T \in \mathcal{T}_h. \end{aligned}$$

For  $v \in H^r(\Omega; \mathbb{R}^d)$ ,  $r > 1/2$ , we define an interpolation  $Iv$  by

$$\int_e Iv \cdot p \, ds = \int_e v \cdot p \, ds, \quad \forall p \in \mathcal{P}_k(e; \mathbb{V}), \quad \forall e \in \mathcal{F}_u, \tag{25}$$

$$\int_T Iv \cdot \phi \, dx = \int_T v \cdot \phi \, dx, \quad \forall \phi \in \mathcal{P}_{k-1}(T; \mathbb{V}), \quad \forall T \in \mathcal{T}_h. \tag{26}$$

Let  $K$  be the  $L^2$  projection from  $L^2(\Omega; \mathbb{K})$  into  $\Gamma_h$ ,

let  $P_i$ ,  $0 \leq i \leq k$ , be the  $L^2$  projection from  $L^2(\Omega; \mathbb{V})$  into

$$\{v \in L^2(\Omega; \mathbb{V}) : v|_T \in \mathcal{P}_i(T; \mathbb{V}), \quad \forall T \in \mathcal{T}_h\}.$$

For these interpolation operators and functions with suitable regularity, we have

$$\|\tau - J\tau\|_{0,h} \leq ch^{\min\{k+1,r\}} \|\tau\|_r, \quad r > 1/2, \tag{27}$$

$$\|v - Iv\|_0 \leq ch^{\min\{k+1,r\}} \|v\|_r, \quad r > 1/2, \tag{28}$$

$$\|\eta - K\eta\|_0 \leq ch^{\min\{k+1,r\}} \|\eta\|_r, \quad r \geq 0, \tag{29}$$

$$\|(v - Iv, 0)\|_h \leq ch^{\min\{k,r-1\}} \|v\|_r, \quad r \geq 1, \tag{30}$$

$$\|v - P_{k-1}v\|_0 \leq ch^{\min\{k,r\}} \|v\|_r, \quad r \geq 0. \tag{31}$$

The estimates (29), (31) are obvious, and the estimates (27), (30), and (28) are consequences of Theorem 3.5 and Theorem 3.4 in [13]. Finally, using the triangle inequality, (29), and (30), one can get

$$\|(v - Iv, \eta - K\eta)\|_h \leq c \left( h^{\min\{k,r-1\}} \|v\|_r + h^{\min\{k+1,\tilde{r}\}} \|\eta\|_{\tilde{r}} \right), \quad r \geq 1, \tilde{r} \geq 0. \tag{32}$$

Optimal convergence rates of the  $L^2$ -errors of  $\sigma$  and  $\gamma$  are obtained by (27), (29), and the following theorem.

**Theorem 2** Suppose that  $(\sigma, u, \boldsymbol{\gamma}), (\sigma_h, u_h, \boldsymbol{\gamma}_h)$  are solutions of (5)–(6) and (16)–(17), respectively. Suppose also that  $\sigma, u$  satisfy (18). Then the following holds:

$$\|\sigma - \sigma_h\|_0 + \|P_{k-1}(u - u_h)\|_0 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0 \leq c(\|\sigma - J\sigma\|_0 + \|\boldsymbol{\gamma} - K\boldsymbol{\gamma}\|_0). \tag{33}$$

Error bounds up to regularity of  $(\sigma, \boldsymbol{\gamma})$  are immediately obtained by (27) and (29).

*Proof* By Lemma 1,  $(\sigma, u, \boldsymbol{\gamma})$  satisfies (16)–(17). The difference of (16)–(17) with  $(\sigma, u, \boldsymbol{\gamma})$  and  $(\sigma_h, u_h, \boldsymbol{\gamma}_h)$  yields

$$(A(\sigma - \sigma_h), \boldsymbol{\tau}) + (u - u_h, \operatorname{div}_h \boldsymbol{\tau}) - \sum_{e \in \mathcal{F}_u} \langle u - u_h, \llbracket \boldsymbol{\tau} n \rrbracket \rangle_e + (\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \boldsymbol{\tau}) = 0, \tag{34}$$

$$(\sigma - \sigma_h, \boldsymbol{\eta} - \operatorname{grad}_h v) + \sum_{e \in \mathcal{F}_\sigma} \langle (\sigma - \sigma_h)n, \llbracket v \rrbracket \rangle_e = 0, \tag{35}$$

for any  $\boldsymbol{\tau} \in \Sigma_h, v \in U_h, \boldsymbol{\eta} \in \Gamma_h$ .

We first prove an auxiliary estimate

$$\|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0 \leq c_0(\|\sigma - J\sigma\|_0 + \|J\sigma - \sigma_h\|_0 + \|\boldsymbol{\gamma} - K\boldsymbol{\gamma}\|_0), \tag{36}$$

with  $c_0 > 0$  independent of mesh sizes. To prove it, let  $P_0(K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h)$  be the  $L^2$  projection of  $K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h$  into the space

$$\mathcal{P}_0(\mathcal{T}_h; \mathbb{K}) := \{\boldsymbol{\eta} \in \Gamma : \boldsymbol{\eta}|_T \in \mathcal{P}_0(T; \mathbb{K}) \forall T \in \mathcal{T}_h\}.$$

It is known (e.g., [6,21]) that there is  $\boldsymbol{\tau}_0 \in H(\operatorname{div}, \Omega; \mathbb{M})$  such that

$$\begin{aligned} \boldsymbol{\tau}_0|_T \in \mathcal{P}_1(T; \mathbb{M}), \quad \forall T \in \mathcal{T}_h, \quad (\boldsymbol{\tau}_0, \boldsymbol{\eta}) &= (P_0(K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h), \boldsymbol{\eta}), \quad \forall \boldsymbol{\eta} \in \mathcal{P}_0(\mathcal{T}_h; \mathbb{K}), \\ \operatorname{div} \boldsymbol{\tau}_0 &= 0, \quad \|\boldsymbol{\tau}_0\|_0 \leq c\|P_0(K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h)\|_0. \end{aligned}$$

By Lemma 9 there is  $\boldsymbol{\tau}_1 \in \Sigma_h \cap H(\operatorname{div}, \Omega; \mathbb{M})$  such that  $\operatorname{div} \boldsymbol{\tau}_1 = 0, \|\boldsymbol{\tau}_1\|_0 \leq c\|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h - \operatorname{skw} \boldsymbol{\tau}_0\|_0$  and  $\operatorname{skw} \boldsymbol{\tau}_1 = K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h - \operatorname{skw} \boldsymbol{\tau}_0$ . Letting  $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_1$ , one can check that  $\operatorname{div} \boldsymbol{\tau} = 0, \llbracket \boldsymbol{\tau} n \rrbracket|_e = 0$  for  $e \in \mathcal{F}_u, \operatorname{skw} \boldsymbol{\tau} = K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h$ , and

$$\|\boldsymbol{\tau}\|_0 \leq \|\boldsymbol{\tau}_0\|_0 + \|\boldsymbol{\tau}_1\|_0 \leq c(\|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0 + \|\boldsymbol{\tau}_0\|_0) \leq c\|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0.$$

If we put this  $\boldsymbol{\tau}$  in (34), then

$$\begin{aligned} 0 &= (A(\sigma - \sigma_h), \boldsymbol{\tau}) + (\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \boldsymbol{\tau}) \\ &= (A(\sigma - \sigma_h), \boldsymbol{\tau}) + (\boldsymbol{\gamma} - K\boldsymbol{\gamma}, \boldsymbol{\tau}) + \|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0^2, \end{aligned} \tag{37}$$

where the second identity is obtained by

$$(\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \boldsymbol{\tau}) = (\boldsymbol{\gamma} - K\boldsymbol{\gamma}, \boldsymbol{\tau}) + (K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \boldsymbol{\tau})$$

combined with  $(K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \operatorname{sym} \boldsymbol{\tau}) = 0$  and  $\operatorname{skw} \boldsymbol{\tau} = K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h$ . The estimate (36) follows by the above equation combined with the Cauchy–Schwarz inequality,  $\|\boldsymbol{\tau}\|_0 \leq c\|K\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0$ , the triangle inequality, and  $\|A\boldsymbol{\tau}\|_0 \leq c\|\boldsymbol{\tau}\|_0$  for  $\boldsymbol{\tau} \in L^2(\Omega; \mathbb{M})$ .

Now we remark two observations. First, the definition of  $J\sigma$  and (35) with  $\boldsymbol{\eta} = 0$  give

$$0 = -(J\sigma - \sigma_h, \operatorname{grad}_h v) + \sum_{e \in \mathcal{F}_\sigma} \langle (J\sigma - \sigma_h)n, \llbracket v \rrbracket \rangle_e, \quad \forall v \in U_h. \tag{38}$$

Second, the definition of  $Iu$  and (34) give

$$(A(\sigma - \sigma_h), \boldsymbol{\tau}) + (Iu - u_h, \operatorname{div}_h \boldsymbol{\tau}) - \sum_{e \in \mathcal{F}_u} \langle Iu - u_h, \llbracket \boldsymbol{\tau} n \rrbracket \rangle_e + (\boldsymbol{\gamma} - \boldsymbol{\gamma}_h, \boldsymbol{\tau}) = 0.$$

Taking  $\tau = J\sigma - \sigma_h$  in this equation, and considering (38) with  $v = Iu - u_h$  and (15), we have

$$(A(\sigma - \sigma_h), J\sigma - \sigma_h) + (\gamma - \gamma_h, J\sigma - \sigma_h) = 0.$$

This is equivalent to

$$\begin{aligned} (A(J\sigma - \sigma_h), J\sigma - \sigma_h) &= -(A(\sigma - J\sigma), J\sigma - \sigma_h) \\ &\quad - (\gamma - K\gamma, J\sigma - \sigma_h) - (K\gamma - \gamma_h, J\sigma - \sigma_h) \\ &= -(A(\sigma - J\sigma), J\sigma - \sigma_h) \\ &\quad - (\gamma - K\gamma, J\sigma - \sigma_h) - (K\gamma - \gamma_h, J\sigma - \sigma), \end{aligned} \tag{39}$$

where the last equality is due to  $(K\gamma - \gamma_h, \sigma_h) = (K\gamma - \gamma_h, \sigma)$  which is obtained from (35) with  $v = 0, \eta = K\gamma - \gamma_h$ . With this observation, the coercivity of  $A$ , the Cauchy–Schwarz inequality, and Young’s inequality, we have

$$\|J\sigma - \sigma_h\|_0^2 \leq C_\epsilon (\|\sigma - J\sigma\|_0^2 + \|\gamma - K\gamma\|_0^2) + \epsilon (\|J\sigma - \sigma_h\|_0^2 + \|K\gamma - \gamma_h\|_0^2),$$

for any  $\epsilon > 0$  with an  $\epsilon$ -dependent constant  $C_\epsilon > 0$ . Using (36) and taking  $\epsilon$  sufficiently small, one can get

$$\|J\sigma - \sigma_h\|_0^2 \leq C'_\epsilon (\|\sigma - J\sigma\|_0^2 + \|\gamma - K\gamma\|_0^2) + \frac{1}{2} \|J\sigma - \sigma_h\|_0^2,$$

with another constant  $C'_\epsilon$ , so  $\|J\sigma - \sigma_h\|_0 \leq c(\|\sigma - J\sigma\|_0 + \|\gamma - K\gamma\|_0)$ . The combination of this and the estimate (36) yields

$$\|J\sigma - \sigma_h\|_0 + \|K\gamma - \gamma_h\|_0 \leq c(\|\sigma - J\sigma\|_0 + \|\gamma - K\gamma\|_0),$$

and the assertion (33) for  $\|\sigma - \sigma_h\|_0$  and  $\|\gamma - \gamma_h\|_0$  follows by the triangle inequality.

To complete the proof, we need to show (33) for  $\|P_{k-1}(u - u_h)\|_0$ . Note that  $d$ -tuples of the Brezzi–Douglas–Marini (BDM) element of order  $k$  is nothing but  $\Sigma_h^{BDM} := \Sigma_h \cap H(\text{div}, \Omega; \mathbb{M})$ , and it is known that there is  $\tau \in \Sigma_h^{BDM}$  such that  $\text{div } \tau = P_{k-1}(u - u_h)$  and  $\|\tau\|_0 \leq c\|P_{k-1}(u - u_h)\|_0$  (see [11, 12]). Inserting this  $\tau$  in (34) we have

$$(A(\sigma - \sigma_h, \tau) + \|P_{k-1}(u - u_h)\|_0^2 + (\gamma - \gamma_h, \tau) = 0.$$

By the Cauchy–Schwarz inequality and  $\|\tau\|_0 \leq c\|P_{k-1}(u - u_h)\|_0$ , one can get

$$\|P_{k-1}(u - u_h)\|_0 \leq c(\|\sigma - \sigma_h\|_0 + \|\gamma - \gamma_h\|_0).$$

Combining this with the previous result of  $\|\sigma - \sigma_h\|_0 + \|\gamma - \gamma_h\|_0$ , the assertion follows.  $\square$

The estimate (33) does not give an optimal  $L^2$ -error bound of  $u$  with respect to the full approximability of  $U_h$ . Now we show that the optimal order of  $L^2$ -error of  $u$  is obtained by the Aubin–Nitsche’s trick when the domain  $\Omega$  satisfies the elliptic regularity assumption, namely, if the triple  $\tilde{\sigma} \in L^2(\Omega; \mathbb{M}), \tilde{u} \in H_0^1(\Omega; \mathbb{V}), \tilde{\gamma} \in L^2(\Omega; \mathbb{K})$  is a solution of the problem

$$(A\tilde{\sigma}, \tau) + (\tilde{\gamma} - \text{grad } \tilde{u}, \tau) = 0, \quad \tau \in L^2(\Omega; \mathbb{M}), \tag{40}$$

$$-(\eta - \text{grad } v, \tilde{\sigma}) = (g, v), \quad (v, \eta) \in H_0^1(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{K}), \tag{41}$$

for  $g \in L^2(\Omega; \mathbb{V})$ , then

$$\|\tilde{\sigma}\|_1 + \|\tilde{u}\|_2 + \|\tilde{\gamma}\|_1 \leq c\|g\|_0 \tag{42}$$

holds with  $c > 0$  depending only on  $\Omega$  and  $A$ . Here  $H_0^1(\Omega; \mathbb{V})$  is the subspace of  $H^1(\Omega; \mathbb{V})$  with vanishing traces [16]. There are several known results on domains satisfying this regularity assumption. For example, it is known that convex polygonal domain in two dimensions [9] and a domain with  $C^2$  boundary in  $\mathbb{R}^d$ ,  $d = 2, 3$ , satisfy the assumption [35, Theorem 7.1].

**Theorem 3** *Suppose that  $(\sigma, u, \gamma)$  and  $(\sigma_h, u_h, \gamma_h)$  are given as in Theorem 2 and  $\Omega$  satisfies the elliptic regularity assumption. Then*

$$\|u - u_h\|_0 \leq ch(\|\sigma - \sigma_h\|_{0,h} + \|(u - u_h, \gamma - \gamma_h)\|_h)$$

holds with  $c > 0$  independent of  $h$ . In particular, assuming that the solution  $(\sigma, u, \gamma)$  is sufficiently regular, this estimate with (24) will give the optimal  $O(h^{k+1})$  error bound for  $\|u - u_h\|_0$ .

*Proof* Let  $(\tilde{\sigma}, \tilde{u}, \tilde{\gamma})$  be the solution of (40)–(41) for  $g = u - u_h$ . Recalling the definition of  $b(\tau; v, \eta)$  in (14), we have

$$(\tilde{\gamma} - \text{grad}_h \tilde{u}, \tau) = b(\tau; \tilde{u}, \tilde{\gamma}), \quad \forall \tau \in \Sigma_h, \tag{43}$$

because  $\tilde{u}$  is in  $H^1(\Omega; \mathbb{V})$ . Since the variational Eq. (41) implies that  $-\text{div } \tilde{\sigma} = u - u_h$ , and  $(\tilde{\sigma}, \gamma - \gamma_h) = 0$  by choosing  $\eta = \gamma - \gamma_h$  and  $v = 0$ , the integration by parts gives

$$\begin{aligned} \|u - u_h\|_0^2 &= (-\text{div } \tilde{\sigma}, u - u_h) - (\tilde{\sigma}, \gamma - \gamma_h) \\ &= (\tilde{\sigma}, \text{grad}_h(u - u_h) - (\gamma - \gamma_h)) - \sum_{e \in \mathcal{F}_\sigma} \langle \tilde{\sigma} n, \llbracket u - u_h \rrbracket \rangle_e \\ &= -b(\tilde{\sigma}; u - u_h, \gamma - \gamma_h). \end{aligned}$$

If we take  $\tau = \sigma - \sigma_h$  in (40), add it with the above equation, and use the identity (43), then we have

$$-a(\tilde{\sigma}, \sigma - \sigma_h) - b(\sigma - \sigma_h; \tilde{u}, \tilde{\gamma}) - b(\tilde{\sigma}; u - u_h, \gamma - \gamma_h) = \|u - u_h\|_0^2.$$

By the Galerkin orthogonality, see (34) and (35), we have

$$\begin{aligned} -a(\tilde{\sigma} - J\tilde{\sigma}, \sigma - \sigma_h) - b(\sigma - \sigma_h; \tilde{u} - I\tilde{u}, \tilde{\gamma} - K\tilde{\gamma}) \\ - b(\tilde{\sigma} - J\tilde{\sigma}; u - u_h, \gamma - \gamma_h) = \|u - u_h\|_0^2. \end{aligned} \tag{44}$$

By (23) and the Cauchy–Schwarz inequality, one can have

$$\begin{aligned} |a(\tilde{\sigma} - J\tilde{\sigma}, \sigma - \sigma_h)| &\leq ch\|\tilde{\sigma}\|_1\|\sigma - \sigma_h\|_0 \leq ch\|\tilde{\sigma}\|_1\|\sigma - \sigma_h\|_{0,h}, \\ |b(\sigma - \sigma_h; \tilde{u} - I\tilde{u}, \tilde{\gamma} - K\tilde{\gamma})| &\leq \|\sigma - \sigma_h\|_{0,h}(\|\tilde{u} - I\tilde{u}, \tilde{\gamma} - K\tilde{\gamma}\|_h) \\ &\leq ch\|\sigma - \sigma_h\|_{0,h}(\|\tilde{\gamma}\|_1 + \|\tilde{u}\|_2), \quad (\text{by (32)}) \\ |b(\tilde{\sigma} - J\tilde{\sigma}; u - u_h, \gamma - \gamma_h)| &\leq \|\tilde{\sigma} - J\tilde{\sigma}\|_{0,h}\|(u - u_h, \gamma - \gamma_h)\|_h \\ &\leq ch\|\tilde{\sigma}\|_1\|(u - u_h, \gamma - \gamma_h)\|_h. \quad (\text{by (27)}) \end{aligned}$$

The conclusion follows by (44), the above three estimates, and (42) with  $g = u - u_h$ .  $\square$

Finally, we remark that the numerical solution  $\sigma_h$  is indeed symmetric. This symmetry follows from the orthogonality between  $\sigma_h$  and  $\Gamma_h$ , and the fact that  $\Sigma_h$  and  $\Gamma_h$  are spaces of polynomials of same degree on each element  $T \in \mathcal{T}_h$ .

### 3.3 Robustness in Nearly Incompressible Materials and Pure Traction Boundary Conditions

In finite element methods for linear elasticity it is known that the energy minimization approach with standard continuous finite elements suffers from the volumetric locking for nearly incompressible materials [8]. When the volumetric locking occurs, the constants in error estimates may grow unboundedly as the incompressibility of the material increases. It is known that mixed methods based on the Hellinger–Reissner formulation of linear elasticity are locking-free [5]. We will prove that our method, which uses a variant of the Hellinger–Reissner formulation, is locking-free as well. For the compliance tensor of the form in (2), nearly incompressible materials have very large  $\lambda$  and the incompressible limit corresponds to  $\lambda = +\infty$ . Therefore, in order to show that a numerical scheme is free from the locking, we need to obtain error estimates such that their involved constants are uniformly bounded for unboundedly growing  $\lambda$ . Note that the error analysis in the proof of Theorem 2 is not enough because the coercivity constant of  $A$  on  $L^2(\Omega; \mathbb{M})$  converges to 0 as  $\lambda \rightarrow +\infty$ . The key idea of our proof is originally from [5] but there is an additional technical difficulty due to the nonconformity of numerical stress with respect to  $H(\text{div}, \Omega; \mathbb{M})$  because the proof in [5] requires that numerical stresses are in  $H(\text{div}, \Omega; \mathbb{M})$ .

In the rest of this section, we assume that compliance tensors have the form in (2) and  $A_D = \partial\Omega$ , as in [5]. Recall that we have extended compliance tensor, originally defined only on symmetric tensors, to be defined on general tensors. This extension is not necessarily unique. However, for simplicity, we will assume that the extended compliance tensor is of the form (2) for  $\boldsymbol{\tau} \in L^2(\Omega; \mathbb{M})$ . Before we state the main result let us introduce preliminary results for the proof. For  $\boldsymbol{\tau} \in L^2(\Omega; \mathbb{M})$  let  $\boldsymbol{\tau}^D := \boldsymbol{\tau} - (1/d) \text{tr}(\boldsymbol{\tau})\mathbf{I}$  with  $d$ , the dimension of Euclidean space  $\mathbb{R}^d$ . The following lemma is proved in [5].

**Lemma 3** For  $\boldsymbol{\tau} \in L^2(\Omega; \mathbb{M})$  and  $A$  of the form in (2), the inequality

$$\|\boldsymbol{\tau}^D\|_0 \leq c\|\boldsymbol{\tau}\|_A \tag{45}$$

holds with  $c$  depending only on  $\mu$  and  $d$ , where  $\|\boldsymbol{\tau}\|_A = (A\boldsymbol{\tau}, \boldsymbol{\tau})^{1/2}$ .

We remark that the interpolation  $I : H^1(\Omega; \mathbb{V}) \rightarrow U_h$  defined by (25) and (26), satisfies

$$-(v - Iv, \text{div}_h \boldsymbol{\tau}) + \sum_{e \in \mathcal{F}_u} \langle v - Iv, \llbracket \boldsymbol{\tau} n \rrbracket \rangle_e = 0, \quad \forall \boldsymbol{\tau} \in \Sigma_h, \tag{46}$$

$$\|v - Iv\|_0 \leq ch\|v\|_1. \tag{47}$$

Here is the main result of the subsection.

**Theorem 4** Suppose that  $A$  is of the form in (2) with positive constants  $\lambda, \mu$ , and suppose also that  $(\boldsymbol{\sigma}, u, \boldsymbol{\gamma}), (\boldsymbol{\sigma}_h, u_h, \boldsymbol{\gamma}_h)$  are solutions of (5)–(6) and (16)–(17), satisfying (18). Then

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|P_{k-1}(u - u_h)\|_0 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0 \leq ch^{\min\{r, k+1\}}(\|\boldsymbol{\sigma}\|_r + \|\boldsymbol{\gamma}\|_r)$$

for  $r > 1/2$ , holds with a constant  $c$  which is uniformly bounded as  $\lambda \rightarrow +\infty$ .

The proof of this theorem is lengthy, so we split it into two lemmas.

**Lemma 4** Under the assumptions same as Theorem 4 we have

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^D\|_0 + \|P_{k-1}(u - u_h)\|_0 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_0 \leq c(\|\boldsymbol{\sigma} - J\boldsymbol{\sigma}\|_0 + \|\boldsymbol{\gamma} - K\boldsymbol{\gamma}\|_0),$$

with  $c$  uniformly bounded as  $\lambda \rightarrow +\infty$ .

*Proof* We begin with deriving a necessary result for the estimate of  $\|(\sigma - \sigma_h)^D\|_0$ . Clearly,  $\|\tau^D\|_0 \leq c\|\tau\|_0$  holds. Then by the triangle inequality and Lemma 3,

$$\begin{aligned} \|(\sigma - \sigma_h)^D\|_0 &\leq \|(\sigma - J\sigma)^D\|_0 + \|(J\sigma - \sigma_h)^D\|_0 \\ &\leq c(\|\sigma - J\sigma\|_0 + \|J\sigma - \sigma_h\|_A). \end{aligned}$$

By (27) there is nothing to do for  $\|\sigma - J\sigma\|_0$  and the desired estimate of  $\|(\sigma - \sigma_h)^D\|_0$  follows from a suitable estimate of  $\|J\sigma - \sigma_h\|_A$ .

To complete the proof we repeat the proof of Theorem 2 with some changes, and we only explain the necessary changes because most of the steps are same. To estimate  $\|J\sigma - \sigma_h\|_A$  we first note that

$$(A\tau, \eta) = (2\mu)^{-1}(\tau, \eta) \quad \text{and} \quad (\tau, \eta) = (\tau^D, \eta), \tag{48}$$

for  $\tau \in L^2(\Omega; \mathbb{M})$  and  $\eta \in L^2(\Omega; \mathbb{K})$ , due to skew-symmetry of  $\eta$ . Note also that  $\|\cdot\|_A \leq c\|\cdot\|_0$  holds with  $c > 0$  independent of  $\lambda$ . We then claim that

$$\|K\gamma - \gamma_h\|_0 \leq c_0(\|\sigma - J\sigma\|_0 + \|J\sigma - \sigma_h\|_A + \|\gamma - K\gamma\|_0), \tag{49}$$

which is an analogue of (36) but with  $c_0$  independent of  $\lambda$ . This estimate follows easily from (37) with the same  $\tau$  by

$$\begin{aligned} \|K\gamma - \gamma_h\|_0^2 &= -(A(\sigma - \sigma_h), \tau) - (\gamma - K\gamma, \tau) \\ &\leq \|\sigma - \sigma_h\|_A \|\tau\|_A + \|\gamma - K\gamma\|_0 \|\tau\|_0 \\ &\leq c(\|\sigma - J\sigma\|_0 + \|J\sigma - \sigma_h\|_A + \|\gamma - K\gamma\|_0) \|K\gamma - \gamma_h\|_0, \end{aligned}$$

in which we used the Cauchy–Schwarz inequality with  $\|\cdot\|_A$  and  $\|\cdot\|_0$ , the triangle inequality, and two facts,  $\|\cdot\|_A \leq c\|\cdot\|_0$  and  $\|\tau\|_0 \leq c\|K\gamma - \gamma_h\|_0$ . The equalities (39) and (48) give

$$\begin{aligned} \|J\sigma - \sigma_h\|_A^2 &= -(A(\sigma - J\sigma), J\sigma - \sigma_h) \\ &\quad - 2\mu(\gamma - K\gamma, A(J\sigma - \sigma_h)) - 2\mu(K\gamma - \gamma_h, A(J\sigma - \sigma)). \end{aligned}$$

Young’s inequality with  $\|\cdot\|_A$  norm gives

$$\|J\sigma - \sigma_h\|_A^2 \leq C_\epsilon (\|\sigma - J\sigma\|_A^2 + \|\gamma - K\gamma\|_A^2) + \epsilon (\|J\sigma - \sigma_h\|_A^2 + \|K\gamma - \gamma_h\|_A^2).$$

Using  $\|\cdot\|_A \leq c\|\cdot\|_0$  and the estimate (49), and taking  $\epsilon$  sufficiently small,

$$\|J\sigma - \sigma_h\|_A^2 \leq C'_\epsilon (\|\sigma - J\sigma\|_0^2 + \|\gamma - K\gamma\|_0^2) + \frac{1}{2} \|J\sigma - \sigma_h\|_A^2.$$

This leads to the desired estimate of  $\|J\sigma - \sigma_h\|_A$ . The rest of the proof is completely analogous to that of Theorem 2, so we omit details.  $\square$

Lemma 4 gives the  $L^2$ -error estimates of  $u$  and  $\gamma$  in Theorem 4. We now prove the estimate of  $\|\sigma - \sigma_h\|_0$  in Theorem 4.

**Lemma 5** *Under the assumptions same as Theorem 4 we have*

$$\|\sigma - \sigma_h\|_0 \leq ch^{\min\{r, k+1\}} (\|\sigma\|_r + \|\gamma\|_r), \quad r > \frac{1}{2},$$

with  $c$  uniformly bounded as  $\lambda \rightarrow +\infty$ .

*Proof* By the triangle inequality  $\|\sigma - \sigma_h\|_0 \leq \|(\sigma - \sigma_h)^D\|_0 + \|\text{tr}(\sigma - \sigma_h)\mathbf{I}\|_0$  and by Lemma 4, it is enough to estimate  $\|\text{tr}(\sigma - \sigma_h)\|_0$ . If we take  $\tau = \mathbf{I}$  in (34), then

$$(A(\sigma - \sigma_h), \mathbf{I}) = \frac{1}{2\mu + d\lambda} \int_{\Omega} \text{tr}(\sigma - \sigma_h) \, dx = 0,$$

so  $\text{tr}(\sigma - \sigma_h)$  is mean-value zero. It is known (see, e.g., [18]) that there is  $w \in H_0^1(\Omega; \mathbb{V})$  such that

$$\text{div}^* w = \text{tr}(\sigma - \sigma_h), \quad \|w\|_1 \leq c \|\text{tr}(\sigma - \sigma_h)\|_0, \tag{50}$$

where  $\text{div}^*$  stands for the column-wise divergence for the  $\mathbb{V}$ -valued function  $w$ . Then

$$\begin{aligned} \|\text{tr}(\sigma - \sigma_h)\|_0^2 &= (\text{tr}(\sigma - \sigma_h), \text{div}^* w) \\ &= (\text{tr}(\sigma - \sigma_h)\mathbf{I}, \text{grad } w) \\ &= d((\sigma - \sigma_h) - (\sigma - \sigma_h)^D, \text{grad } w). \end{aligned} \tag{51}$$

By the Cauchy–Schwarz inequality and the inequality in (50),

$$((\sigma - \sigma_h)^D, \text{grad } w) \leq c \|(\sigma - \sigma_h)^D\|_0 \|\text{tr}(\sigma - \sigma_h)\|_0. \tag{52}$$

We now claim that

$$(\sigma - \sigma_h, \text{grad } w) \leq ch^{\min\{r, k+1\}} \|\sigma\|_r \|\text{tr}(\sigma - \sigma_h)\|_0, \quad r > \frac{1}{2}, \tag{53}$$

and the proof will be completed by combining this, (52), and Lemma 4.

To prove (53), notice first that we can obtain the following from (35) using the integration by parts and taking  $\eta = \mathbf{0}$ :

$$(\text{div}_h(\sigma - \sigma_h), v) - \sum_{e \in \mathcal{F}_u} \langle \llbracket (\sigma - \sigma_h)n \rrbracket, v \rangle_e = 0, \quad \forall v \in U_h. \tag{54}$$

By the integration by parts and (54),

$$\begin{aligned} (\sigma - \sigma_h, \text{grad } w) &= -(\text{div}_h(\sigma - \sigma_h), w) + \sum_{e \in \mathcal{F}_u} \langle \llbracket (\sigma - \sigma_h)n \rrbracket, w \rangle_e \\ &= -(\text{div}_h(\sigma - \sigma_h), w - Iw) + \sum_{e \in \mathcal{F}_u} \langle \llbracket (\sigma - \sigma_h)n \rrbracket, w - Iw \rangle_e. \end{aligned}$$

Let  $\sigma'_h \in \Sigma_h^{BDM}$  be the object obtained from  $\sigma$  by the canonical interpolation of the BDM element of order  $k$ . Then, in the last form of the above,  $\sigma_h$  can be replaced by  $\sigma'_h$  due to (46), and we get

$$(\sigma - \sigma_h, \text{grad } w) = (\text{div}_h(\sigma - \sigma'_h), w - Iw). \tag{55}$$

Here the jump terms disappeared because  $\sigma'_h$  has no normal jump. By the integration by parts,

$$\begin{aligned} (\text{div}_h(\sigma - \sigma'_h), w - Iw) &= -((\sigma - \sigma'_h), \text{grad}_h(w - Iw)) \\ &\quad + \sum_{e \in \mathcal{F}_\sigma} \langle (\sigma - \sigma'_h)n, \llbracket w - Iw \rrbracket \rangle_e \\ &\leq \|\sigma - \sigma'_h\|_{0,h} \|(w - Iw, 0)\|_h \\ &\leq ch^{\min\{r, k+1\}} \|\sigma\|_r \|w\|_1, \end{aligned}$$



where the last inequality follows by an approximation estimate of BDM interpolation [11] and (30). Now (53) follows by (55), the above inequality, and (50).  $\square$

As the final topic in this section, we discuss necessary modifications of our numerical scheme for the problems with pure traction boundary conditions. For these boundary conditions the space of rigid body motions on  $\Omega$ , denoted by  $RM(\Omega)$ , is the kernel of the problem, so it needs to be ruled out in the finite element spaces. More precisely, we define

$$\widehat{\Sigma}_h = \{\boldsymbol{\tau} \in \Sigma_h : \boldsymbol{\tau}n = 0 \text{ on } \partial\Omega\},$$

$$\widehat{U}_h = \left\{ v \in U_h : \int_{\Omega} v \cdot w \, dx = 0 \text{ for } w \in RM(\Omega) \right\}.$$

It is easy to see that  $\|\cdot\|_{0,h}$  in (12) is a norm on  $\widehat{\Sigma}_h$ . To see that  $\|(\cdot, \cdot)\|_h$  in (13) is a norm on  $\widehat{U}_h \times \widehat{\Gamma}_h$ , we only need to check that  $\|(v, \boldsymbol{\eta})\|_h = 0$  for  $(v, \boldsymbol{\eta}) \in \widehat{U}_h \times \widehat{\Gamma}_h$  implies that  $v = \boldsymbol{\eta} = 0$ . If  $\|(v, \boldsymbol{\eta})\|_h = 0$ , then  $(\boldsymbol{\eta} - \text{grad}_h v)|_T = 0$  for all  $T \in \mathcal{T}_h$ , and thus  $\text{sym grad}_h v = 0$ . In other words,  $v|_T$  is a rigid body motion on each  $T \in \mathcal{T}_h$ . Since  $\llbracket v \rrbracket_e = 0$  for all  $e \in \mathcal{F}_\sigma$  due to  $\|(v, \boldsymbol{\eta})\|_h = 0$ ,  $v$  is a rigid body motion on  $\Omega$  and therefore  $v = 0$  by the definition of  $\widehat{U}_h$ . Then  $\boldsymbol{\eta} = 0$  holds because  $\|(0, \boldsymbol{\eta})\|_h = 0$ .

### 4 Numerical Results

For our staggered DG method, we test a simple model problem in two dimensions with the exact solution,

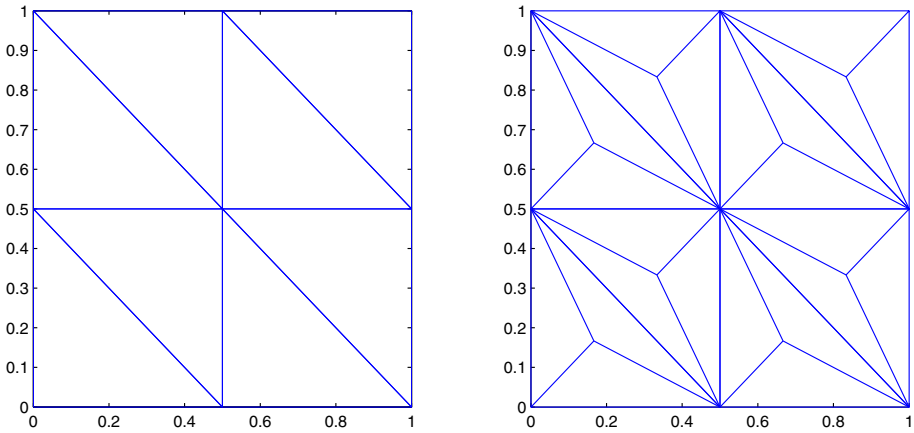
$$u(x, y) = \begin{pmatrix} \sin(\pi x) & \sin(\pi y) \\ x(1-x) & y(1-y) \end{pmatrix}, \quad \forall (x, y) \in \Omega, \tag{56}$$

where the computational domain  $\Omega$  is the unit square,  $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ .

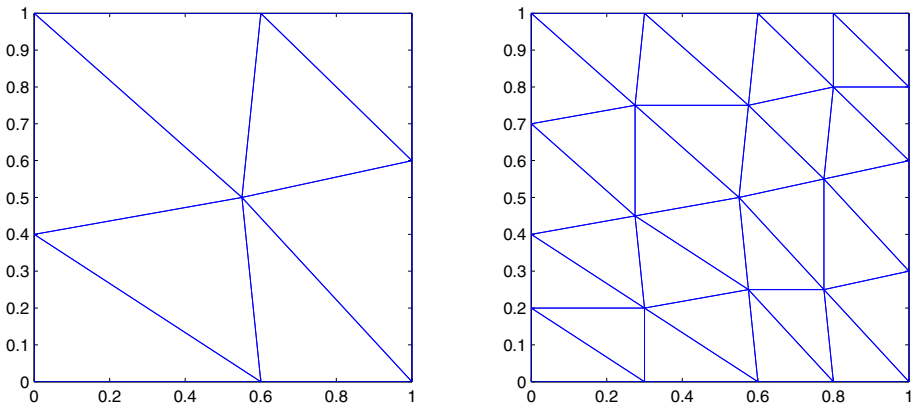
The computational domain is divided into a conforming triangulation  $\mathcal{M}_h$  and then each triangle  $M$  in  $\mathcal{M}_h$  is subdivided into three small triangles to form the resulting triangulation  $\mathcal{T}_h$ . All the finite element functions are piecewise polynomials in the triangulation  $\mathcal{T}_h$  with partial continuity across the edges of the triangulation as described in Sect. 3.1. The order of polynomials of approximation spaces is taken to be  $k = 1$  in the following experiments for simplicity.

In our experiment we test our algorithm for both structured and unstructured meshes by decreasing the mesh size. We consider the exact solution in (56) with the material parameters  $\mu = 1$  and  $\lambda = 1$ . In the structured mesh, the computational domain is divided into  $N \times N$  uniform squares and then each square is divided into two triangles to obtain the triangulation  $\mathcal{M}_h$ . The triangles  $M$  in  $\mathcal{M}_h$  are then divided into three subtriangles to form the resulting triangulation  $\mathcal{T}_h$ , see Fig. 2 for an example. In the unstructured mesh, the initial triangulation  $\mathcal{M}_{h_0}$  is refined into  $\mathcal{M}_{h_1}$  by connecting the mid points of edges of each triangles  $M$  in  $\mathcal{M}_{h_0}$ , see Fig. 3. The refinement is done successively to get more refined mesh up to the level 5,  $\mathcal{M}_{h_5}$ . Similarly as in Fig. 2, each triangles in  $\mathcal{M}_{h_l}$  are then subdivided into three small triangles to form the resulting triangulation  $\mathcal{T}_{h_l}$ . In the following experiments, the Dirichlet boundary condition is imposed on the whole boundary of the computational domain.

In Table 1, errors are presented by increasing  $N$ , i.e., by decreasing the mesh size  $h = 1/N$ . We use the notations  $E(u) := u - u_h$ ,  $E(\boldsymbol{\sigma}) := \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ ,  $E(\boldsymbol{\eta}) := \boldsymbol{\eta} - \boldsymbol{\eta}_h$ , and  $E(u, \boldsymbol{\eta}) := (u - u_h, \boldsymbol{\eta} - \boldsymbol{\eta}_h)$ . We observe that  $L^2$ -errors for  $u - u_h$ ,  $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ , and  $\boldsymbol{\eta} - \boldsymbol{\eta}_h$  follow the order 2 and the error  $\|(u - u_h, \boldsymbol{\eta} - \boldsymbol{\eta}_h)\|_h$  follows the order 1, which are optimal with respect to the given polynomial order  $k = 1$ .



**Fig. 2** Structured mesh: triangulation  $\mathcal{M}_h$  (left) and  $\mathcal{T}_h$  (right) when  $N = 2$



**Fig. 3** Unstructured mesh: level 0 mesh,  $\mathcal{M}_{h_0}$  (left) and level 1 mesh,  $\mathcal{M}_{h_1}$  (right)

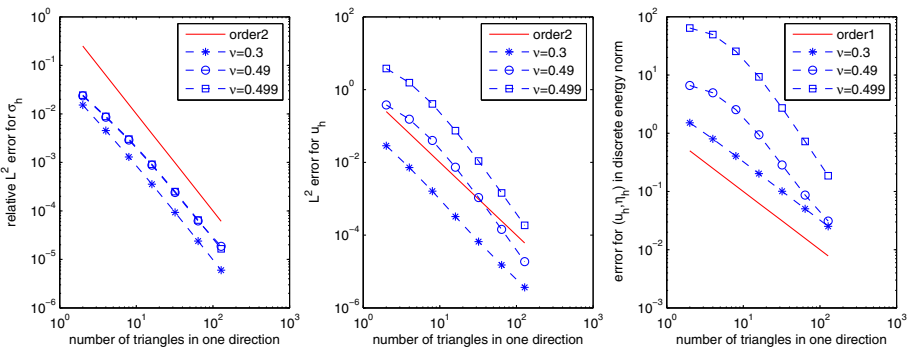
**Table 1** Structured mesh: convergence of errors for the approximation using  $P_1$  elements,  $E(u) := u - u_h$ ,  $E(\sigma) := \sigma - \sigma_h$ ,  $E(\eta) := \eta - \eta_h$ , and  $E(u, \eta) := (u - u_h, \eta - \eta_h)$

$N$	$\ E(u)\ _0$	Order	$\ E(\sigma)\ _0$	Order	$\ E(\eta)\ _0$	Order	$\ E(u, \eta)\ _h$	Order
2	2.66e-02	–	5.62e-01	–	1.71e-01	–	1.50e-00	–
4	6.00e-03	2.14	1.60e-01	1.81	6.74e-02	1.34	7.94e-01	0.91
8	1.30e-03	2.20	4.35e-02	1.87	2.42e-02	1.47	4.02e-01	0.98
16	2.71e-04	2.26	1.15e-02	1.91	7.80e-03	1.63	2.01e-01	1.00
32	6.11e-05	2.14	2.90e-03	1.98	2.20e-03	1.82	1.00e-01	1.00
64	1.46e-05	2.06	7.42e-04	1.96	5.74e-04	1.93	5.03e-02	0.99

In Table 2, errors are presented by increasing the refinement level  $l$  in the unstructured meshes. We also observe that  $L^2$ -errors for  $u - u_{h_l}$ ,  $\sigma - \sigma_{h_l}$ , and  $\eta - \eta_{h_l}$  follow the order  $k + 1$  and the errors for  $(u - u_{h_l}, \eta - \eta_{h_l})$  in the discrete energy norm follow the order  $k$ .

**Table 2** Unstructured mesh: convergence of errors for the approximation using  $P_1$  elements,  $E(u) := u - u_{h_l}$ ,  $E(\sigma) := \sigma - \sigma_{h_l}$ ,  $E(\eta) := \eta - \eta_{h_l}$ , and  $E(u, \eta) := (u - u_{h_l}, \eta - \eta_{h_l})$

Level $l$	$\ E(u)\ _0$	Order	$\ E(\sigma)\ _0$	Order	$\ E(\eta)\ _0$	Order	$\ E(u, \eta)\ _{h_l}$	Order
0	2.55e-02	–	6.06e-01	–	2.03e-01	–	1.36e-00	–
1	5.40e-03	2.25	1.61e-01	1.90	6.97e-02	1.54	7.44e-01	0.87
2	1.10e-03	2.26	4.32e-02	1.89	2.31e-02	1.58	3.75e-01	0.98
3	2.43e-04	2.18	1.14e-02	1.92	7.39e-03	1.64	1.87e-01	1.00
4	5.65e-05	2.10	2.95e-03	1.94	2.16e-03	1.76	9.38e-02	1.00
5	1.36e-05	2.04	7.54e-04	1.97	5.87e-04	1.88	4.68e-02	1.00

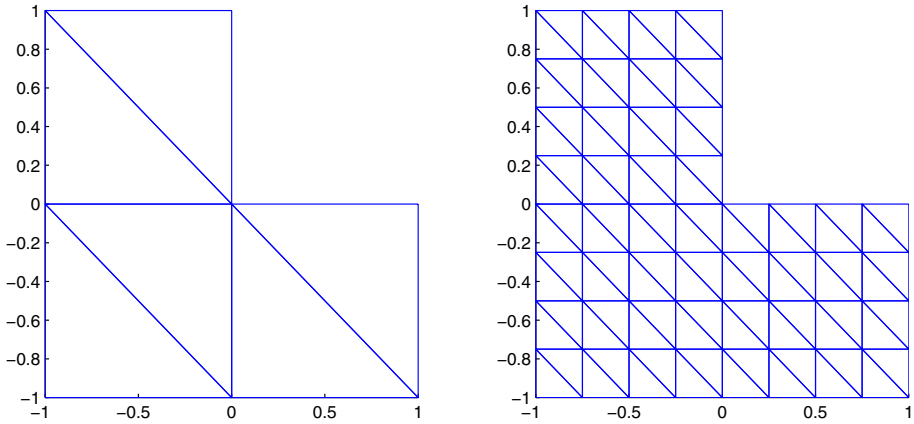


**Fig. 4** Plot of errors with respect to the number of *triangles* in one direction as  $\nu$  approaches the limit value 0.5

To show locking-free property of our numerical scheme, we consider a model problem with the same exact solution in (56) and with the shear modulus  $\mu = 1$  and the material parameter  $\lambda$ , which is determined by the given Poisson ratio  $\nu$  and equation  $\lambda = 2\mu\nu/(1 - 2\nu)$ . We compare errors for the various values of  $\nu = 0.3, 0.49, 0.499$  and the errors are plotted in Fig. 4 with respect to the number of triangles in one direction. For the stress tensor, the relative  $L^2$ -error is calculated by using  $\|\sigma_h - \sigma\|_0/\|f\|_0$ , where  $f$  is the right hand side of (1) for the exact solution in (56) with the given  $\mu$  and  $\lambda$ , see also (2). The uniform meshes as in Fig. 2 are used in our experiment. The relative  $L^2$ -errors for the stress tensor follow the growth of the second order to the size of meshes and they seem to approach the same value as  $\nu$  gets closer to the limit value 0.5. For  $L^2$ -errors of displacement, as the mesh size gets closer to zero, we can observe more reduction in the ratio of errors when  $\nu$  approaches the limit value 0.5. In other words, the slope of errors is getting steeper for decreasing the mesh size  $h$  as  $\nu$  gets closer to 0.5. Though the  $\|(\cdot, \cdot)\|_h$ -norm for  $(u_h, \eta_h)$  is not covered by our theory, we can numerically observe the locking-free property for that case (Fig. 4).

We now consider an example with a singular solution in a non-convex domain. The computational domain is an L-shaped domain  $\Omega = [-1, 1] \times [-1, 1] \setminus (0, 1) \times (0, 1) \subset \mathbb{R}^2$  (see Fig. 5), where a model problem with the following exact solution is considered

$$u(x, y) = (x^2 + y^2)^\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \alpha \geq 0.$$



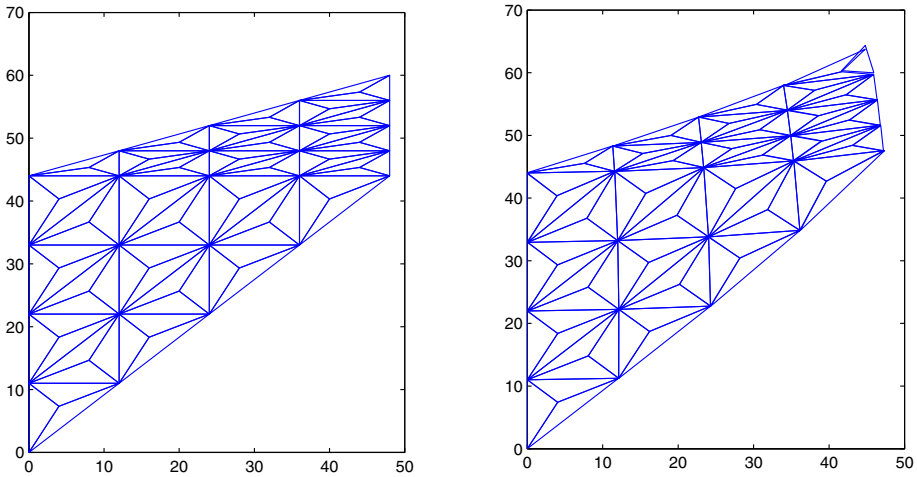
**Fig. 5** L-shaped domain: initial mesh  $\mathcal{M}_{h_0}$  (left) and level2 mesh  $\mathcal{M}_{h_2}$  (right)

**Table 3** L-shaped domain with a singular solution: convergence of errors for the approximation using  $P_1$  elements,  $E(u) := u - u_{h_l}$ ,  $E(\sigma) := \sigma - \sigma_{h_l}$ ,  $E(\eta) := \eta - \eta_{h_l}$ , and  $E(u, \eta) := (u - u_{h_l}, \eta - \eta_{h_l})$

Level $l$	$\ E(u)\ _0$	Order	$\ E(\sigma)\ _0$	Order	$\ E(\eta)\ _0$	Order	$\ E(u, \eta)\ _{h_l}$	Order
0	1.15e-01	–	1.45e-00	–	4.51e-01	–	1.01e-00	–
1	5.12e-02	1.77	8.93e-02	0.70	4.03e-01	0.16	8.53e-01	0.25
2	1.89e-02	1.43	5.55e-02	0.68	2.75e-01	0.55	5.87e-01	0.54
3	6.64e-03	1.51	3.48e-02	0.67	1.75e-01	0.64	3.80e-01	0.63
4	2.35e-03	1.49	2.19e-02	0.66	1.10e-01	0.66	2.42e-01	0.65
5	8.44e-04	1.46	1.38e-02	0.66	6.98e-02	0.66	1.53e-01	0.66

Here  $\alpha$  is chosen as  $1/3$ . With this choice of  $\alpha$ ,  $u$  is in  $H^r(\Omega; \mathbb{V})$  for any  $0 \leq r < 2\alpha + 1 (\simeq 1.67)$ . Starting from the initial uniform mesh  $\mathcal{M}_{h_0}$ , the refined meshes  $\mathcal{M}_{h_l}$  are obtained as in Fig. 5. Numerical results are presented in Table 3 and again here the Dirichlet boundary condition is enforced on the whole boundary of the L-shaped domain. From the results, we observe that the errors except the  $L^2$  errors of  $u$  follow the order about 0.66 determined by the regularity of the model problem. In addition, the  $L^2$ -errors of  $u$  are observed to be about 1.4 which is affected by the limited regularity of the L-shaped domain.

As the last numerical experiment, we consider Cook’s membrane problem with the geometry in Fig. 6. The left part of the boundary is clammed on the wall, a surface load in vertical direction is enforced on the opposite right part of the boundary, and the zero surface load on the rest part of the boundary. The body force  $f$  is zero and the Young modulus  $E = 2900$  and Poisson ratio  $\nu = 0.3$  are given. The deformed mesh from the Cook’s membrane problem is plotted in Fig. 6.



**Fig. 6** Cook’s membrane problem: staggered mesh on the geometry (left) and deformed mesh (right)

**Acknowledgments** We are indebted to the anonymous referees for their helpful comments which significantly improved the presentation and quality of this paper.

### Appendix

In this appendix we provide some lemmas which are used in the proof of our main results.

Let  $\mathcal{M}_h$  be a shape regular triangulation of a bounded domain  $\Omega$ .

For simplicity, we use  $\|\cdot\|_r$  for norms instead of  $\|\cdot\|_{r,M}$  when it is clear that functions are clearly defined on  $M$ .

**Lemma 6** Let  $T_i, i = 1, \dots, d + 1$ , be subsimplices of  $M \in \mathcal{M}_h$  obtained by the barycentric subdivision. For a given  $k \geq 1$ , let

$$V_{M,k} = \{v \in H^1(M; \mathbb{R}^d) : v|_{T_i} \in \mathcal{P}_k(T_i; \mathbb{R}^d) \text{ for } 1 \leq i \leq d + 1, v|_{\partial M} = 0\}, \tag{57}$$

$$Q_{M,k} = \left\{ p \in L^2(M) : p|_{T_i} \in \mathcal{P}_k(T_i), \int_M p \, dx = 0 \right\}. \tag{58}$$

Then there exists  $\beta > 0$  independent of  $M \in \mathcal{M}_h$  such that

$$\inf_{0 \neq p \in Q_{M,k}} \sup_{0 \neq w \in V_{M,k+1}} \frac{(\operatorname{div} w, p)}{\|w\|_1 \|p\|_0} \geq \beta. \tag{59}$$

*Proof* Note that (59) is the inf-sup condition of mixed finite element method  $(V_{M,k+1}, Q_{M,k})$  for the Stokes equation, so we simply refer to appropriate literature. When  $d = 2$  with  $k \geq 3$ , (59) is obtained from the stability of Scott–Vogelius elements for the Stokes equation [29]. For  $d = 2$  with  $k = 1, 2$ , it is reported in [27]. When  $d = 3$ , it is proved in [37].  $\square$

**Lemma 7** Suppose  $V_{M,k}$  and  $Q_{M,k}$  are defined as in (57) and (58). Then there exists  $c > 0$  which only depends on the shape regularity of  $\mathcal{M}_h$  and  $k \geq 1$  such that, for any given  $p \in Q_{M,k}$ , there exists  $w \in V_{M,k+1}$  satisfying  $\operatorname{div} w = p$  and  $\|w\|_1 \leq c \|p\|_0$ .

*Proof* Let  $M \in \mathcal{M}_h$  be fixed. Note that  $\operatorname{div} w \in Q_{M,k}$  for  $w \in V_{M,k+1}$ , so we can regard  $\operatorname{div}$  as a linear map from  $V_{M,k+1}$  to  $Q_{M,k}$ . This  $\operatorname{div}$  is surjective because, if not, then there exists  $0 \neq p \in Q_{M,k}$  such that  $(\operatorname{div} w, p) = 0$  for all  $w \in V_{M,k+1}$ , but it contradicts (59). Suppose that  $\{p_i\}_{i=1}^m$  is an orthonormal basis of  $Q_{M,k}$  in  $L^2(M)$  and define

$$\alpha_i := \inf_{w \in V_{M,k+1}, \operatorname{div} w = p_i} \frac{(\operatorname{div} w, p_i)}{\|w\|_1}, \quad \alpha_M := \min_{1 \leq i \leq m} \alpha_i. \tag{60}$$

By finite dimensionality of the polynomial spaces  $V_{M,k+1}$  and  $Q_{M,k}$ , there exists  $w_i \in V_{M,k+1}$  such that  $\operatorname{div} w_i = p_i$  and  $\|w_i\|_1 = \alpha_i^{-1}$ . Suppose that a given  $0 \neq p \in Q_{M,k}$  can be expressed as  $p = \sum_{i=1}^m c_i p_i$  with coefficients  $c_i \in \mathbb{R}$ ,  $1 \leq i \leq m$ . Then  $\|p\|_0^2 = \sum_{i=1}^m c_i^2$ . If we take  $w = \sum_{i=1}^m c_i w_i$ , then  $\operatorname{div} w = p$  and

$$\|w\|_1 \leq \sum_{i=1}^m |c_i| \|w_i\|_1 = \sum_{i=1}^m \frac{|c_i|}{\alpha_i} \leq \left( \sum_{i=1}^m c_i^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^m \alpha_i^{-2} \right)^{\frac{1}{2}} \leq \|p\|_0 \alpha_M^{-1} \sqrt{m},$$

by the triangle inequality, the fact  $\|w_i\|_1 = \alpha_i^{-1}$  for  $1 \leq i \leq m$ , the Cauchy-Schwarz inequality, and the definition of  $\alpha_M$  in (60). To complete proof, note that the  $\alpha_i$ 's in (60) are independent of the diameter of  $M$  by standard scaling argument. By shape regularity assumption of  $\mathcal{M}_h$  and standard compactness argument (see, e.g., [31, Lemma3.1]), there exists  $\alpha > 0$  depending only on the shape regularity of  $\mathcal{M}_h$  and  $k$  such that  $\alpha \leq \alpha_M$  for any  $M \in \mathcal{M}_h$ . The proof is completed.  $\square$

Let  $\mathbb{L} = \mathbb{V}$  if  $d = 2$  and  $\mathbb{L} = \mathbb{M}$  if  $d = 3$ . For given  $k \geq 1$ , we define  $\mathcal{E}_{M,0}$  as

$$\mathcal{E}_{M,0} = \{ \xi \in H^1(M; \mathbb{L}) : \xi|_{T_i} \in \mathcal{P}_{k+1}(T_i; \mathbb{L}), 1 \leq i \leq d + 1, \xi|_{\partial M} = 0 \}. \tag{61}$$

We also define  $S$  and  $\chi$  as

$$S \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \frac{1}{2} (\xi_1 \ \xi_2) \quad \text{for } \xi \in \mathcal{E}_{M,0}, \quad \chi(r) = \begin{pmatrix} 0 & r \\ -r & 0 \end{pmatrix} \quad \text{for } r \in \mathbb{R} \quad \text{if } d = 2,$$

$$S\xi = \frac{1}{2} (\xi^T - (\operatorname{tr} \xi) \mathbf{I}) \quad \text{for } \xi \in \mathcal{E}_{M,0}, \quad \chi \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} \quad \text{if } d = 3,$$

where  $\xi^T$  is the transpose of  $\xi$  and  $\mathbf{I}$  is the identity  $3 \times 3$  matrix. Note that  $S$  and  $\chi$  are algebraic isomorphisms. One can verify by a direct computation that

$$\operatorname{skw} \operatorname{curl} \xi = \chi \operatorname{div} S\xi, \quad \xi \in \mathcal{E}_{M,0}. \tag{62}$$

**Lemma 8** Assume that  $\mathcal{E}_{M,0}$  is defined as in (61) for some  $k \geq 1$ . If  $\zeta \in \mathcal{E}_{M,0}$ , then  $(\operatorname{curl} \zeta)n|_{\partial M} = 0$  for the unit normal vector field  $n$  on  $\partial M$ .

*Proof* If  $d = 2$ , then  $\operatorname{curl} \zeta$  is the  $\pi/2$ -rotation of  $\operatorname{grad} \zeta$ . Denoting by  $t$ , the tangent vector obtained by  $\pi/2$ -rotation of the normal vector  $n$  on  $\partial M$ ,  $(\operatorname{grad} \zeta)t = 0$  on  $\partial M$  because  $\zeta$  is vanishing on  $\partial M$ . Then we can see that  $(\operatorname{curl} \zeta)n = (\operatorname{grad} \zeta)t = 0$  on  $\partial M$ .

When  $d = 3$ , let  $\zeta_i, i = 1, 2, 3$ , be the  $i$ -th row of  $\zeta$ . It suffices to show that  $\text{curl } \zeta_i \cdot n \equiv 0$  on  $\partial M$ . By the Stokes' theorem, for any  $\phi \in C^1(\overline{M})$ ,

$$\begin{aligned} 0 &= \int_M \text{div}(\text{curl } \zeta_i)\phi \, dx = \int_{\partial M} (\text{curl } \zeta_i \cdot n)\phi \, ds - \int_M (\text{curl } \zeta_i) \cdot \text{grad } \phi \, dx \\ &= \int_{\partial M} (\text{curl } \zeta_i \cdot n)\phi \, ds + \int_{\partial M} (\zeta_i \times n) \cdot \text{grad } \phi \, ds - \int_M \zeta_i \cdot \text{curl}(\text{grad } \phi) \, dx \\ &= \int_{\partial M} (\text{curl } \zeta_i \cdot n)\phi \, ds + \int_{\partial M} (\zeta_i \times n) \cdot \text{grad } \phi \, ds. \end{aligned}$$

Since  $\zeta_i \equiv 0$  on  $\partial M$ ,  $\int_{\partial M} (\zeta_i \times n) \cdot \text{grad } \phi \, ds = 0$  and then  $\text{curl } \zeta_i \cdot n \equiv 0$  on  $\partial M$  because  $\phi \in C^1(\overline{M})$  is arbitrary. □

**Lemma 9** For a given  $k \geq 1$ , define  $\Gamma_{M,0}$  as

$$\Gamma_{M,0} = \left\{ \eta \in L^2(M; \mathbb{K}) : \eta|_{T_i} \in \mathcal{P}_k(T_i; \mathbb{K}), 1 \leq i \leq d + 1, \int_M \eta \, dx = 0 \right\},$$

and recall the definition of  $\Sigma_{M,0}$  in (21) with same  $k$ . For any  $\eta \in \Gamma_{M,0}$  there is  $\tau \in \Sigma_{M,0}$  such that  $\text{div } \tau = 0$ ,  $\text{skw } \tau = \eta$ , and  $\|\tau\|_0 \leq c\|\eta\|_0$  with  $c > 0$  independent of  $M \in \mathcal{M}_h$ .

*Proof* We only prove the three dimensional case because the two dimensional one can be proved similarly. For the given  $k$ , we consider  $V_{M,k+1}, Q_{M,k}$  as in (57), (58), and  $\mathcal{E}_{M,0}$  as in (61). By the definition of  $\chi$ , note that  $\chi^{-1}\eta \in L^2(M; \mathbb{V})$  for  $\eta \in \Gamma_{M,0}$  and each component of  $\chi^{-1}\eta$  is in  $Q_{M,k}$ . Applying Lemma 7 to each component of  $\chi^{-1}\eta$ , one can find  $\xi \in H^1(M; \mathbb{M})$  such that  $\text{div } \xi = \chi^{-1}\eta, \|\xi\|_1 \leq c\|\eta\|_0$ , and each row can be identified with an element in  $V_{M,k+1}$ . If we set  $\tau = \text{curl } S^{-1}\xi$ , then  $\text{div } \tau = 0$  and also  $\tau \in \Sigma_{M,0}$  by Lemma 8 because  $S^{-1}\xi \in \mathcal{E}_{M,0}$ . Furthermore, by (62) and the assumptions on  $\xi$ ,

$$\begin{aligned} \text{skw } \tau &= \text{skw } \text{curl } S^{-1}\xi = \chi \text{div } \xi = \eta, \\ \|\tau\|_0 &= \|\text{curl } S^{-1}\xi\|_0 \leq c\|\xi\|_1 \leq c\|\eta\|_0. \end{aligned}$$

The proof is completed. □

## References

1. Amara, M., Thomas, J.M.: Equilibrium finite elements for the linear elastic problem. *Numer. Math.* **33**(4), 367–383 (1979). doi:[10.1007/BF01399320](https://doi.org/10.1007/BF01399320)
2. Arnold, D.N., Awanou, G., Winther, R.: Finite elements for symmetric tensors in three dimensions. *Math. Comp.* **77**(263), 1229–1251 (2008). doi:[10.1090/S0025-5718-08-02071-1](https://doi.org/10.1090/S0025-5718-08-02071-1)
3. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39**(5), 1749–1779 (2001/02). doi:[10.1137/S0036142901384162](https://doi.org/10.1137/S0036142901384162)
4. Arnold, D.N., Brezzi, F., Douglas, J.J.: PEERS: a new mixed finite element for plane elasticity. *J. Appl. Math. Jpn.* **1**, 347–367 (1984). doi:[10.1016/j.jfa.2010.02.007](https://doi.org/10.1016/j.jfa.2010.02.007)
5. Arnold, D.N., Douglas, J.J., Gupta, C.P.: A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.* **45**(1), 1–22 (1984). doi:[10.1007/BF01379659](https://doi.org/10.1007/BF01379659)
6. Arnold, D.N., Falk, R.S., Winther, R.: Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Comp.* **76**(260), 1699–1723 (2007). doi:[10.1090/S0025-5718-07-01998-9](https://doi.org/10.1090/S0025-5718-07-01998-9). (electronic)
7. Arnold, D.N., Winther, R.: Mixed finite elements for elasticity. *Numer. Math.* **92**(3), 401–419 (2002). doi:[10.1007/s002110100348](https://doi.org/10.1007/s002110100348)
8. Babuška, I., Suri, M.: Locking effects in the finite element approximation of elasticity problems. *Numer. Math.* **62**(4), 439–463 (1992). doi:[10.1007/BF01396238](https://doi.org/10.1007/BF01396238)

9. Bacuta, C., Bramble, J.H.: Regularity estimates for solutions of the equations of linear elasticity in convex plane polygonal domains. *Z. Angew. Math. Phys.* **54**(5), 874–878 (2003). doi:[10.1007/s00033-003-3211-4](https://doi.org/10.1007/s00033-003-3211-4). Special issue dedicated to Lawrence E. Payne
10. Boffi, D., Brezzi, F., Fortin, M.: Reduced symmetry elements in linear elasticity. *Commun. Pure Appl. Anal.* **8**(1), 95–121 (2009). doi:[10.3934/cpaa.2009.8.95](https://doi.org/10.3934/cpaa.2009.8.95)
11. Brezzi, F., Douglas, J.J., Marini, L.D.: Two families of mixed finite elements for second order elliptic problems. *Numer. Math.* **47**(2), 217–235 (1985). doi:[10.1007/BF01389710](https://doi.org/10.1007/BF01389710)
12. Brezzi, F., Fortin, M.: *Mixed and Hybrid Finite Element Methods*, Springer Series in Computational Mathematics, vol. 15. Springer, New York (1992)
13. Chung, E.T., Engquist, B.: Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions. *SIAM J. Numer. Anal.* **47**(5), 3820–3848 (2009). doi:[10.1137/080729062](https://doi.org/10.1137/080729062)
14. Cockburn, B., Gopalakrishnan, J., Guzmán, J.: A new elasticity element made for enforcing weak stress symmetry. *Math. Comp.* **79**(271), 1331–1349 (2010). doi:[10.1090/S0025-5718-10-02343-4](https://doi.org/10.1090/S0025-5718-10-02343-4)
15. Cockburn, B., Shi, K.: Superconvergent HDG methods for linear elasticity with weakly symmetric stresses. *IMA J. Numer. Anal.* **33**(3), 747–770 (2013). doi:[10.1093/imanum/drs020](https://doi.org/10.1093/imanum/drs020)
16. Evans, L.C.: *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (1998)
17. Farhloul, M., Fortin, M.: Dual hybrid methods for the elasticity and the Stokes problems: a unified approach. *Numer. Math.* **76**(4), 419–440 (1997). doi:[10.1007/s002110050270](https://doi.org/10.1007/s002110050270)
18. Girault, V., Raviart, P.A.: *Finite element methods for Navier-Stokes equations: Theory and algorithms*, Springer Series in Computational Mathematics, vol. 5. Springer-Verlag, Berlin (1986). doi:[10.1007/978-3-642-61623-5](https://doi.org/10.1007/978-3-642-61623-5)
19. Gopalakrishnan, J., Guzmán, J.: A second elasticity element using the matrix bubble. *IMA J. Numer. Anal.* **32**(1), 352–372 (2012). doi:[10.1093/imanum/drq047](https://doi.org/10.1093/imanum/drq047)
20. Gurtin, M.E.: *An introduction to continuum mechanics*, Mathematics in Science and Engineering, vol. 158. Academic Press Inc.(Harcourt Brace Jovanovich Publishers), New York (1981)
21. Guzmán, J.: A unified analysis of several mixed methods for elasticity with weak stress symmetry. *J. Sci. Comput.* **44**(2), 156–169 (2010). doi:[10.1007/s10915-010-9373-2](https://doi.org/10.1007/s10915-010-9373-2)
22. Guzmán, J., Neilan, M.: Symmetric and conforming mixed finite elements for plane elasticity using rational bubble functions. *Numer. Math.* **126**(1), 153–171 (2014). doi:[10.1007/s00211-013-0557-1](https://doi.org/10.1007/s00211-013-0557-1)
23. Johnson, C., Mercier, B.: Some equilibrium finite element methods for two-dimensional elasticity problems. *Numer. Math.* **30**(1), 103–116 (1978)
24. Kim, H.H., Chung, E.T., Lee, C.S.: A staggered discontinuous Galerkin method for the Stokes system. *SIAM J. Numer. Anal.* **51**(6), 3327–3350 (2013). doi:[10.1137/120896037](https://doi.org/10.1137/120896037)
25. Kim, K.Y.: Analysis of some low-order nonconforming mixed finite elements for linear elasticity problem. *Numer. Methods. Partial. Differ. Equ.* **22**(3), 638–660 (2006). doi:[10.1002/num.20114](https://doi.org/10.1002/num.20114)
26. Morley, M.E.: A family of mixed finite elements for linear elasticity. *Numer. Math.* **55**(6), 633–666 (1989). doi:[10.1007/BF01389334](https://doi.org/10.1007/BF01389334)
27. Qin, J.: On the convergence of some low order mixed finite elements for incompressible fluids. ProQuest LLC, Ann Arbor, MI (1994). Thesis (Ph.D.)-The Pennsylvania State University
28. Qiu, W., Demkowicz, L.: Mixed  $hp$ -finite element method for linear elasticity with weakly imposed symmetry: stability analysis. *SIAM J. Numer. Anal.* **49**(2), 619–641 (2011). doi:[10.1137/100797539](https://doi.org/10.1137/100797539)
29. Scott, L.R., Vogelius, M.: Conforming finite element methods for incompressible and nearly incompressible continua. In: *Large-scale computations in fluid mechanics, Part 2* (La Jolla, Calif., 1983), Lectures in Appl. Math., vol. 22, pp. 221–244. Amer. Math. Soc., Providence, RI (1985)
30. Soon, S.C., Cockburn, B., Stolarski, H.K.: A hybridizable discontinuous Galerkin method for linear elasticity. *Int. J. Numer. Methods Eng.* **80**(8), 1058–1092 (2009). doi:[10.1002/nme.2646](https://doi.org/10.1002/nme.2646)
31. Stenberg, R.: Analysis of mixed finite elements methods for the Stokes problem: a unified approach. *Math. Comp.* **42**(165), 9–23 (1984). doi:[10.2307/2007557](https://doi.org/10.2307/2007557)
32. Stenberg, R.: A family of mixed finite elements for the elasticity problem. *Numer. Math.* **53**(5), 513–538 (1988). doi:[10.1007/BF01397550](https://doi.org/10.1007/BF01397550)
33. Stenberg, R.: Two low-order mixed methods for the elasticity problem. In: *The mathematics of finite elements and applications*. VI (Uxbridge, 1987), pp. 271–280. Academic Press, London (1988)
34. Stenberg, R.: Weakly symmetric mixed finite elements for linear elasticity. In: Abdulle, A., Deparis, S., Kressner, D., Nobile, F., Picasso, M. (eds.) *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, Lecture Notes in Computational Science and Engineering, vol. 103, pp. 3–18. Springer International Publishing (2015). doi:[10.1007/978-3-319-10705-9\\_1](https://doi.org/10.1007/978-3-319-10705-9_1)
35. Valent, T.: *Boundary Value Problems of Finite Elasticity*, Springer Tracts in Natural Philosophy, vol. 31. Springer, New York (1988). doi:[10.1007/978-1-4612-3736-5](https://doi.org/10.1007/978-1-4612-3736-5). Local theorems on existence, uniqueness, and analytic dependence on data



36. de Veubeke, B.M.F.: Stress function approach. In: Proceedings of the World Congress on Finite Element Methods in Structural Mechanics, vol. 5, pp. J.1–J.51 (1975)
37. Zhang, S.: A new family of stable mixed finite elements for the 3D Stokes equations. *Math. Comp.* **74**(250), 543–554 (2005). doi:[10.1090/S0025-5718-04-01711-9](https://doi.org/10.1090/S0025-5718-04-01711-9)