SYMMETRIC NON-CONFORMING MIXED FINITE ELEMENTS FOR LINEAR ELASTICITY

J. GOPALAKRISHNAN AND J. GUZMÁN

ABSTRACT. We present a family of mixed methods for linear elasticity, that yield exactly symmetric, but only weakly conforming, stress approximations. The method is presented in both two and three dimensions (on triangular and tetrahedral meshes). The method is efficiently implementable by hybridization. The degrees of freedom of the Lagrange multipliers, which approximate the displacements at the faces, solve a symmetric positive-definite system. The design and analysis of this method is motivated by a new set of unisolvent degrees of freedom for symmetric polynomial matrices. These new degrees of freedom are also used to give a new simple calculation of the dimension of the space of polynomial symmetric matrix fields with vanishing normal traces and zero divergence on a tetrahedron. Such a dimension count was important in the development of the symmetric $H(\mathrm{div})$ conforming methods found in [*Finite elements for symmetric tensors in three dimensions*, Math. Comp. 77 (2008), no. 263, 1229–1251].

1. INTRODUCTION

Mixed finite element methods are popular methods in solid mechanics since they avoid locking and provide direct approximations to stresses. However, stress finite elements needed for mixed methods are difficult to design due to two requirements. First, due to the conservation of angular momentum, the stress tensor is required to be symmetric. Second, the forces on a mesh face shared by two mesh elements must be in equilibrium, i.e., the stress approximation must lie in $H(\mathrm{div})$.

Stress finite elements satisfying both these requirements have been designed in [1, 3, 4, 8]. But the main drawback of these methods is that they cannot be implemented as efficiently as other mixed methods. Their stress elements have too many degrees of freedom and techniques like hybridization usually available for mixed methods are not available for these methods. A related family of methods use composite elements and yield symmetric approximations to the stresses [6, 24]. However, they use piecewise discontinuous polynomials within each element.

Recognizing the difficulties caused by imposing both the above mentioned requirements, researchers have pursued design of methods that relax one of the above two requirements. The first category of such methods weakly impose stress symmetry, while maintaining exact $H(\mathrm{div})$-conformity. These methods introduce a Lagrange multiplier approximating the non-symmetric part of the displacement gradient while enforcing stress symmetry weakly [2, 7, 5, 13, 17, 19, 20, 21, 25, 26, 28].

One can equally well consider methods that relax the other requirement, namely the $H(\mathrm{div})$-conformity of stress, yielding a second category of methods. The contributions of [9, 10, 22, 23, 29, 30] fit into this category, and so does the contribution of this paper. These methods yield non-conforming, but symmetric, stress approximations. It is not clear if the relaxation of the first or the second requirements is better, and research into both avenues are actively pursued in the mathematical literature. The methods with either relaxations, either in the first or the second category, can be more efficiently implemented compared to the above mentioned strongly symmetric conforming methods.

To describe our contribution, let us introduce the linear elasticity problem

$$\mathrm{div}\,\sigma = f \quad \text{in } \Omega, \tag{1.1a}$$

$$\mathcal{A}\sigma - \varepsilon(u) = 0 \quad \text{in } \Omega, \tag{1.1b}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1.1c}$$

where $\Omega \subset \mathbb{R}^d$ (for $d = 2, 3$). Here, displacement is represented by the function $u : \Omega \mapsto \mathbb{R}^d$. The stress is denoted by $\sigma : \Omega \mapsto \mathbb{S}$, where $\mathbb{S}$ denotes the set of real symmetric $d \times d$ matrices. The compliance tensor is denoted by $\mathcal{A}(x)$, and is assumed to be a bounded, symmetric, positive definite tensor over $\mathbb{S}$. The given load is denoted by the vector function $f : \Omega \mapsto \mathbb{R}^d$. The linearized strain tensor is $\varepsilon(u) = (\mathrm{grad}\,u + (\mathrm{grad}\,u)')/2$. Here and throughout, transposes are denoted by $'$, and differential operators are applied row-wise.

Our finite element spaces for the stresses and displacements are very simple. Let Ω_h be a simplicial mesh of Ω (satisfying the standard finite element conformity assumptions). Our spaces, for $k \geq 1$, are given by

$$V^h := \{\tau : \tau|_K \in \mathcal{P}^{k+1}(K, \mathbb{S}), \text{ for all } K \in \Omega_h, \text{ and the moments of } \tau n$$
$$\text{up to degree } k \text{ are continuous across element interfaces }\}, \tag{1.2a}$$

$$W^h := \{w : w|_K \in \mathcal{P}^k(K, \mathbb{R}^d), \text{ for all } K \in \Omega_h\}. \tag{1.2b}$$

Here $\mathcal{P}^k(D, \mathbb{R}^d)$ (or $\mathcal{P}^k(D, \mathbb{S})$) denotes the space of functions from $D \mapsto \mathbb{R}^d$ (or $D \mapsto \mathbb{S}$, resp.) whose scalar components are polynomials of degree at most k. Above, n denotes a normal vector on the interelement boundaries. If τn were continuous across element interfaces, then τ would be in $H(\mathrm{div}, \Omega, \mathbb{S}) := \{\eta \in L^2(\Omega, \mathbb{S}) : \mathrm{div}\,\eta \in L^2(\Omega, \mathbb{R}^d)\}$, the space where the exact stresses lie. Clearly, functions τ in V^h are, in general, *not* in $H(\mathrm{div}, \Omega, \mathbb{S})$, so our method is non-conforming.

Our method finds $(\sigma^h, u^h) \in V^h \times W^h$ by satisfying

$$(\mathcal{A}\sigma^h, \tau)_\Omega + (u^h, \mathrm{div}_h\,\tau)_\Omega = 0 \tag{1.3a}$$

$$(\mathrm{div}_h\,\sigma^h, \omega)_\Omega = (f, \omega)_\Omega \tag{1.3b}$$

for all $(\tau, \omega) \in V^h \times W^h$. Note that above $\sigma^h \notin H(\mathrm{div}, \Omega, \mathbb{S})$ in general, and $\mathrm{div}_h\,(\cdot)$ denotes the element-wise divergence. Here and throughout, for vector functions $\zeta, \theta : D \mapsto \mathbb{R}^d$, the notation $(\zeta, \theta)_D$ denotes the integral over D of their dot product, while for matrix functions ζ and $\theta : D \mapsto \mathbb{S}$, the same notation $(\zeta, \theta)_D$ denotes the integral over D of their Frobenius inner product $\zeta : \theta \equiv \mathrm{tr}(\zeta'\theta)$. Our new mixed nonconforming method is recommended by the simplicity of its elements, and the

possibility of hybridization for efficient implementations. The main drawback is that we are only able to prove suboptimal convergence rates.

The key to our analysis is a new set of unisolvent degrees of freedom for the local space $\mathcal{P}^{k+1}(K, \mathbb{S})$ for any $k \geq 1$ and space dimensions $d = 2, 3$. The degrees of freedom of the space $\mathcal{P}^{k+1}(K, \mathbb{S})$ use edge moments in three dimensions and vertex values in two-dimensions. However, let us emphasize at the outset that these edge or vertex degrees of freedom, unlike the strongly symmetric conforming methods for simplicial meshes, will not make our method inefficient. In fact, we will present the hybrid form of the method in Section 4, where we introduce Lagrange multipliers of degree k that approximate the displacements on mesh faces. We will show that the other variables can be eliminated locally to obtain a final symmetric-positive definite system for only the degrees of freedom associated with the Lagrange multipliers.

Nevertheless, we also investigate if one can construct a reduced space V^h which does not contain edge moments in three dimensions (or vertex values in two-dimensions) as degrees of freedom. This is indeed possible as shown in Section 5. The reduced space replaces $\mathcal{P}^{k+1}(K, \mathbb{S})$ in (1.2a) with a space $V(K)$ satisfying $\mathcal{P}^k(K, \mathbb{S}) \subset V(K) \subset \mathcal{P}^{k+1}(K, \mathbb{S})$. The method using the reduced space has the same convergence rates as the corresponding full space. Although the dimension of $V(K)$ is smaller than the dimension of $\mathcal{P}^{k+1}(K, \mathbb{S})$, the simplicity of the full polynomial space $\mathcal{P}^{k+1}(K, \mathbb{S})$ might be more attractive.

As far as we know, all the other non-conforming mixed elements to date [9, 10, 22, 23, 29, 30] are low-order elements. In contrast, we develop a family of simplicial elements (one for each $k \geq 1$) in both two and three dimensions. All but the elements in [9] are rectangular elements. Arnold and Winther in [9], define two simplicial elements in two dimensions only. Their first element uses the displacement space W^h above with $k = 1$ and their stress space has exactly the same dimension as our reduced stress space for $k = 1$. In fact, the degrees of freedom of our reduced space corresponding to $k = 1$ is the same as their first stress space. Their second element [9] uses piecewise rigid displacements as the space of displacements and hence the displacement error can be at most first-order accurate.

The new degrees of freedom for $\mathcal{P}^{k+1}(K, \mathbb{S})$, while playing a key role in the design and analysis of our non-conforming mixed method, also find other applications. As an interesting application unrelated to our method, we give an alternative proof of the dimension count of the space

$$\mathcal{M}^k := \{\tau \in \mathcal{P}^k(K, \mathbb{S}) : \operatorname{div} \tau = 0 \text{ and } \tau n|_{\partial K} = 0\}. \qquad (1.4)$$

This space is essential in the development of conforming mixed methods [4, 8]. Its dimension count in particular is important in proving the unisolvency of the degrees of freedom. The proof of the dimension count of the space $\mathcal{M}^k$ in three dimensions given in [4] is substantially more involved than the two-dimensional proof [8]. The alternative proof we give here is much simpler.

The paper is organized as follows. In the next section we give the new unisolvent degrees of freedom for the space $\mathcal{P}^{k+1}(K, \mathbb{S})$. As corollaries, we prove the dimension count of $\mathcal{M}^k$ and the well-posedness of the mixed method (1.3) with spaces (1.2b), (1.2a). In Section 3 we provide the error analysis of the mixed method. In Section 4 we present the hybrid form of the mixed method. Finally, in Section (5) we present the reduced element.

2. A set of degrees of freedom for symmetric polynomial matrices

In this section we give a unisolvent set of degrees of freedom (d.o.f) for $\mathcal{P}^{k+1}(K, \mathbb{S})$. Here K is a tetrahedra in three dimensions $(d = 3)$, or a triangle in two dimensions $(d = 2)$.

The unisolvent set of d.o.f, in the three-dimensional $(d = 3)$ case, is the following set of linear functionals:

$$\ell_\rho(\sigma) = (\sigma, \rho)_K, \qquad \text{for all } \rho \in \mathcal{P}^{k-1}(K, \mathbb{S}), \tag{2.5a}$$

$$\ell_\mu(\sigma) = \langle \sigma n, \mu \rangle_F, \qquad \text{for all } \mu \in \mathcal{P}^k(F, \mathbb{R}^d), \text{ for all faces } F \text{ of } K, \tag{2.5b}$$

$$\ell_s(\sigma) = \langle \sigma n^- \cdot n^+, s \rangle_e, \quad \text{for all } s \in \mathcal{P}^{k+1}(e, \mathbb{R}), \text{ for all edges } e \text{ of } K. \tag{2.5c}$$

Here, for each edge e, n^+ and n^- are the normal vectors of the two faces that intersect at the edge e. Note that $\sigma n^- \cdot n^+ = \sigma n^+ \cdot n^-$ since σ is symmetric. The notation $\langle a, b \rangle_S$ denotes integration of the (scalar) product of a and b over (lower dimensional) S. In the case of two dimensions $(d = 2)$, the degrees of freedom are given by

$$\ell_\rho(\sigma) = (\sigma, \rho)_K, \qquad \text{for all } \rho \in \mathcal{P}^{k-1}(K, \mathbb{S}), \tag{2.5a$'$}$$

$$\ell_\mu(\sigma) = \langle \sigma n, \mu \rangle_F, \qquad \text{for all } \mu \in \mathcal{P}^k(F, \mathbb{R}^d), \text{ for all edges } F \text{ of } K, \tag{2.5b$'$}$$

$$\ell_x(\sigma) = \sigma(x) n^- \cdot n^+, \qquad \text{for all vertices } x \text{ of } K. \tag{2.5c$'$}$$

For each vertex x of K the vectors n^+ and n^- are the normal vectors of the two edges that intersect at x. The main result of this section is the following theorem. In the $k = 0$ case, we omit (2.5a) and (2.5a$'$) from the set of d.o.f.

THEOREM 2.1 (Unisolvency): *Assume* $k \geq 0$. *Let* K *be a triangle* $(d = 2)$ *or a tetrahedra* $(d = 3)$. *Any* σ *in* $\mathcal{P}^{k+1}(K, \mathbb{S})$ *is uniquely determined by the degrees of freedom given by* (2.5) *for* $d = 3$ *or* (2.5$'$) *for* $d = 2$.

Proof. We prove the result in two steps. First, we show that the number of d.o.f equal

$$\dim(\mathcal{P}^{k+1}(K, \mathbb{S})) = \frac{d(d+1)}{2} \frac{(k+1+d)!}{(k+1)! d!}.$$

Consider $d = 3$ case. The number of functionals in (2.5a), (2.5b), and (2.5c) are

$$6 \frac{k(k+1)(k+2)}{6}, \quad 4 \frac{3(k+1)(k+2)}{2}, \quad \text{and } 6(k+2),$$

respectively. These add up to $6(k+2)(k+3)(k+4) = \dim(\mathcal{P}^{k+1}(K, \mathbb{S}))$. The $d = 2$ case is similar.

The second step is to show that if all the d.o.f applied to $\sigma \in \mathcal{P}^{k+1}(K, \mathbb{S})$ vanish, then σ vanishes. We only consider the $d = 3$ case as the $d = 2$ case is similar. Accordingly, let F_i and F_j be any two distinct faces of a tetrahedron K. We denote by λ_i the barycentric coordinate of K that vanishes on F_i, and by n_i a normal on F_i.

We first claim that $\sigma n_i \cdot n_j$ vanishes on F_i and F_j. Clearly, by (2.5c), it vanishes on the edge shared by F_i and F_j. Hence, on F_i, $\sigma n_i \cdot n_j|_{F_i} = \lambda_j|_{F_i} v$ for some v in $\mathcal{P}^k(F_i, \mathbb{R})$. But by (2.5a), $\langle \sigma n_i, n_j v \rangle_{F_i} = \langle \lambda_j v, v \rangle_{F_i} = 0$, so $v \equiv 0$. Thus $\sigma n_i \cdot n_j$ vanishes on F_i. A similar argument on F_j shows that it vanishes there as well. Consequently,

$$\sigma n_i \cdot n_j = \lambda_i \lambda_j w \tag{2.6}$$

for some w in $\mathcal{P}^{k-1}(K,\mathbb{R})$. But, by (2.5a), we have $(\sigma, n_i n_j' w)_K = (\lambda_i \lambda_j w, w)_K = 0$, which implies $w \equiv 0$. Thus $\sigma n_i \cdot n_j \equiv 0$ on K.

Any vector in $\mathbb{R}^3$ can be expressed in terms of any three of the normals $\{n_i\}$ since they form a basis of $\mathbb{R}^3$ (the fourth normal can be written as a linear combination of the other three). Since $\sigma n_i \cdot n_j$ vanishes for any distinct i and j, we find that $\sigma(x)$ is a symmetric matrix whose quadratic form vanishes (at every point x in K). Hence σ vanishes on K.

Note that in the $k = 0$ case, the argument leading to (2.6) already shows that σ vanishes and there is no need to use (2.5a). $\qquad\square$

We conclude this section with two corollaries. The first provides a considerable simplification of the proof of a previously known result [4, Theorem 7.2] concerning $\mathcal{M}^k$, a space fundamental in the analysis of known stress elements for conforming mixed methods in linear elasticity [4, 8] (see (1.4) for its definition).

COROLLARY 2.2: *For any non-negative integer k,*

$$\dim(\mathcal{M}^k) = \begin{cases} 0 & \text{if } k \le 3 \\ (k-2)(k-3)/2 & \text{if } d = 2 \text{ and } k \ge 4, \\ (k+2)(k-2)(k-3)/2 & \text{if } d = 3 \text{ and } k \ge 4. \end{cases}$$

Proof. We only prove the $d = 3$ case, as the $d = 2$ case is similar. Let $\mathcal{R}_1$ denote the space of rigid displacements of the form $a + b \times x$ for some $a, b \in \mathbb{R}^3$. Let $\mathcal{R}_\perp^k$ denote the orthogonal complement of $\mathcal{R}_1$ in $\mathcal{P}^k(K,\mathbb{R}^d)$. Let $k \ge 1$. Then we have the decomposition

$$\mathcal{P}^{k-1}(K,\mathbb{S}) = \varepsilon(\mathcal{R}_\perp^k) \oplus \mathcal{S}^{k-1} \tag{2.7}$$

where $\mathcal{S}^{k-1}$ is the orthogonal complement of $\varepsilon(\mathcal{R}_\perp^k)$ in $\mathcal{P}^{k-1}(K,\mathbb{S})$. Decomposing ρ in (2.5a) as $\rho = \varepsilon(r) + \eta$, we find that we may split all the interior d.o.f in (2.5a) into two subcategories:

$$\ell_{\varepsilon(r)}(\sigma) = (\sigma, \varepsilon(r))_K, \qquad\qquad \text{for all } r \in \mathcal{R}_\perp^k, \text{ and} \qquad (2.5a\varepsilon)$$

$$\ell_\eta = (\sigma, \eta)_K, \qquad\qquad \text{for all } \eta \in \mathcal{S}^{k-1}. \qquad (2.5a\eta)$$

Now, we claim that

$$\mathcal{M}^{k+1} = \{\sigma \in \mathcal{P}^{k+1}(K,\mathbb{S}): \text{the d.o.f in } (2.5a\varepsilon), (2.5b), \text{ and } (2.5c),$$
$$\text{applied to } \sigma \text{ vanish}\}. \tag{2.8}$$

Since the $\subseteq$–containment is easy, we only show the reverse. Let F_i be an arbitrary face. Then, the zero face and edge d.o.f ((2.5b) and (2.5c)) imply, by the same argument that led to (2.6), that $\sigma n_i \cdot n_m|_{F_i} = 0$ for all three indices $m \ne i$. Thus $\sigma n_i|_{F_i} = 0$, so σn vanishes on ∂K. That $\operatorname{div}\sigma$ also vanishes follows from $(2.5a\varepsilon)$ and integration by parts. Hence σ is in $\mathcal{M}^{k+1}$.

Next, the d.o.f in (2.8) are linearly independent functionals as they form a subset of a unisolvent set of d.o.f (by Theorem 2.1). Hence $\dim(\mathcal{M}^{k+1})$ can be obtained by simply subtracting their number from $\dim(\mathcal{P}^{k+1}(K,\mathbb{S}))$

$$\dim(\mathcal{M}^{k+1}) = \dim(\mathcal{P}^{k+1}(K,\mathbb{S})) - \dim(\mathcal{R}_\perp^k) - 4\frac{3(k+1)(k+2)}{2} - 6(k+2). \tag{2.9}$$

Since $\dim(\mathcal{R}_\perp^k) = 3\dim(\mathcal{P}^k(K,\mathbb{R})) - 6$, upon simplifying, we find this equals $(k + 3)(k-1)(k-2)/2$ for $k + 1 \geq 4$. The right hand side of (2.9) is zero for $k + 1 = 3$, showing that $\mathcal{M}^3$ is trivial. Since $\mathcal{M}^k \subseteq \mathcal{M}^3$ for $k < 3$, the spaces $\mathcal{M}^0$, $\mathcal{M}^1$, and $\mathcal{M}^2$ are also trivial. $\qquad\square$

Remark 2.3. The proof of the corollary also reveals the following procedure for constructing a basis for $\mathcal{M}^{k+1}$: First, by the Gram-Schmidt process, find an orthonormal bases for the two subspaces in (2.7). Using these bases, split the d.o.fs in (2.5a) into the d.o.fs in (2.5aε) and (2.5aη). Having done this split, consider the shape functions that form a dual basis to all the d.o.fs in (2.5). From this collection, if we omit the shape functions associated to the d.o.fs in (2.5aε), (2.5b), and (2.5c), then we obtain a basis for $\mathcal{M}^{k+1}$. This is because of the characterization (2.8) that we just proved.

In the next corollary, we prove that the mixed method (1.3) is well defined.

COROLLARY 2.4: *The mixed method* (1.3) *using the spaces* V^h *and* W^h *from* (1.2) *has a unique solution.*

Proof. Since (1.3) is a square linear system it is enough to prove uniqueness. To this end, we set $f = 0$. Then, by (1.3b), $\mathrm{div}_h\, \sigma^h = 0$. So, putting $\tau = \sigma^h$ in (1.3a) we find $(\mathcal{A}\sigma^h, \sigma^h)_\Omega = 0$. Thus $\sigma^h = 0$.

In order to show that u_h also vanishes, we use integration by parts in (1.3a) to get

$$\sum_{K \in \Omega_h} -(\varepsilon(u^h), \tau)_K + \langle u^h, \tau n\rangle_{\partial K} = 0 \quad \text{for all } \tau \in V^h. \tag{2.10}$$

On each K we define $\tau|_K \in \mathcal{P}^{k+1}(K, \mathbb{S})$ by setting the d.o.f (in (2.5) or (2.5$'$)) as follows: Set $\ell_\rho(\tau) \equiv (\tau, \rho)_K = -(\varepsilon(u^h), \rho)_K$, for all $\rho \in \mathcal{P}^{k-1}(K, \mathbb{S})$, and let all other degrees of freedom vanish. It is easy to see that this τ is in V^h. Using it in (2.10),

$$\sum_{K \in \Omega_h} (\varepsilon(u^h),\, \varepsilon(u^h))_K = 0.$$

Thus, $\varepsilon(u^h|_K) = 0$ on each $K \in \Omega_h$.

Next, we show that u^h is continuous and vanishes on $\partial\Omega$. For this, on any interior face F shared by two mesh elements K_+ and K_-, we let $u_\pm^h$ denote the trace of u^h on F from $K_\pm$. Also denote the unit outward normals on $K_\pm$ by $n_\pm$. Then the jump of u^h is denoted by $[\![u^h n']\!] = u_+^h n_+' + u_-^h n_-'$. On faces $F \subseteq \partial\Omega$, $[\![u^h n']\!] = u^h n'$. Now, for any $K \in \Omega_h$, define $\tau|_K \in \mathcal{P}^{k+1}(K, \mathbb{S})$ by setting $\ell_\mu(\tau) \equiv \langle \tau n, \mu\rangle_F = \langle [\![u^h n']\!]n, \mu\rangle_F$ for all $F \subseteq \partial K$ and letting all the remaining degrees of freedom vanish. The composite function τ is then clearly in V^h. Using this τ in (2.10), and recalling that $\varepsilon(u^h|_K) = 0$, we find that $\langle [\![u^h n']\!], [\![u^h n']\!]\rangle_F = 0$ for all mesh faces F (or all mesh edges in the $d = 2$ case). Thus u^h is continuous and vanishes on the boundary. Combined with the fact that $\varepsilon(u^h) = 0$, this implies that u^h vanishes on Ω. $\qquad\square$

Remark 2.5. The local element d.o.fs in (2.5)$'$ furnish a set of global d.o.fs, given a triangulation (considering, say the $d = 2$ case). On an edge shared by two elements, the edge d.o.fs in (2.5b)$'$ coalesce into one global d.o.f suitable for enforcing the weak continuity. But note that the vertex d.o.fs in (2.5c)$'$ do not coalesce, i.e., to each mesh vertex, we must associate as many global d.o.fs as the number of edges connected to that vertex.

3. Error Estimates

In this section we prove the following error estimates for the mixed method (1.3). We assume throughout that the mesh is shape regular and use h to denote the maximum of the diameters of all mesh elements.

THEOREM 3.1: *Suppose $k \geq 1$ and $\sigma \in H^1(\Omega, \mathbb{S})$. Then, for any $1 \leq r \leq k$,*

$$\|\sigma - \sigma^h\|_{L^2(\Omega)} + \|u - u^h\|_{L^2(\Omega)} \leq Ch^r \|u\|_{H^{r+1}(\Omega)}, \tag{3.11}$$

Moreover, if full elliptic regularity holds (see (3.29)), then

$$\|u - u^h\|_{L^2(\Omega)} \leq Ch^{r+1} \|u\|_{H^{r+1}(\Omega)}. \tag{3.12}$$

In the remainder of this section, we prove Theorem 3.1 in several steps. Above and throughout, we have adopted the usual convention of denoting by C a positive constant independent of h, whose specific value at different occurrences may vary.

3.1. **Error estimate for the stress variable.** Error estimation of the stress begins with the error equations, obtained by subtracting the discrete equations from the exact ones:

$$(\mathcal{A}(\sigma - \sigma^h), \tau)_\Omega + (u - u^h, \mathrm{div}_h \, \tau)_\Omega = \langle u, \tau n \rangle_{\partial \Omega_h} \tag{3.13a}$$

$$(\mathrm{div}_h \, (\sigma - \sigma^h), \omega)_\Omega = 0 \tag{3.13b}$$

for all $(\tau, \omega) \in V^h \times W^h$, where

$$\langle u, \tau n \rangle_{\partial \Omega_h} = \sum_{K \in \Omega_h} \langle u, \tau n \rangle_{\partial K}. \tag{3.14}$$

This last term in (3.13a) term measures the *consistency error*. It is in general non-zero because V^h is not $H(\mathrm{div})$-conforming.

We need a projector into the finite element space with a commutativity property. To reveal the main idea, we proceed by assuming the existence of such a projector and verify the assumption later (in § 3.2).

Assumption 3.2. There is a subspace $D_\Pi \subseteq H(\mathrm{div}, \Omega, \mathbb{S})$ and a continuous projector $\Pi : D_\Pi \mapsto V^h$ such that

$$\mathrm{div}_h \, (\Pi \sigma) = P \mathrm{div}_h \, \sigma \qquad \text{for all } \sigma \in D_\Pi, \tag{3.15}$$

where $P : L^2(\Omega, \mathbb{R}^d) \mapsto W^h$ is the L^2-orthogonal projection.

THEOREM 3.3: *If Π is as in Assumption 3.2 and $k \geq 1$, then whenever $\sigma \in D_\Pi$,*

$$C\|\sigma - \sigma^h\|_{L^2(\Omega)} \leq \|\sigma - \Pi\sigma\|_{L^2(\Omega)} + \inf_{w^h \in \mathcal{W}^h} \|\varepsilon(u - w^h)\|_{L^2(\Omega)}, \tag{3.16}$$

where $\mathcal{W}^h$ is the set of all continuous functions in W^h that vanish on $\partial \Omega$.

Proof. We first note that using (3.15) and using (3.13b) we have that

$$\mathrm{div}_h \, (\Pi \sigma - \sigma^h) = 0. \tag{3.17}$$

Hence, setting $\tau = \Pi\sigma - \sigma^h$ in (3.13a) and rearranging, we find that

$$(\mathcal{A}(\Pi\sigma - \sigma^h), \Pi\sigma - \sigma^h)_\Omega = (\mathcal{A}(\Pi\sigma - \sigma), \Pi\sigma - \sigma^h)_\Omega + \langle u, (\Pi\sigma - \sigma^h)n \rangle_{\partial \Omega_h}.$$

Now, since $\Pi\sigma - \sigma^h$ is in V^h, we can replace u by $u - w^h$ in the last term, for any w^h in $\mathcal{W}^h$. Furthermore, integrating by parts on each element K, and using (3.17), we find $\langle u - w^h, (\Pi\sigma - \sigma^h)n\rangle_{\partial K} = (\varepsilon(u - w^h), \Pi\sigma - \sigma^h)_K$. Thus, by Cauchy-Schwarz inequality,

$$C\|\Pi\sigma - \sigma^h\|_{L^2(\Omega)} \leq \|\Pi\sigma - \sigma\|_{L^2(\Omega)} + \|\varepsilon(u - w^h)\|_{L^2(\Omega)} \tag{3.18}$$

where we have also used that $\mathcal{A}$ is positive definite. The theorem now follows from triangle inequality. $\qquad\square$

Remark 3.4. To measure errors in the momentum balance one may additionally consider the error terms e_1, e_2 defined by

$$e_1^2 = \sum_{K \in \Omega_h} h_K \|[\![(\sigma - \sigma_h)n]\!]\|_{L^2(\partial K)}^2, \qquad e_2 = \mathrm{div}_h\,(\sigma - \sigma_h).$$

Estimate (3.18) implies a bound on e_1 via an inverse estimate:

$$e_1 \leq \left(\sum_{K \in \Omega_h} h_K \|[\![(\sigma - \Pi\sigma)n]\!]\|_{L^2(\partial K)}^2 \right)^{1/2} + C\|\sigma - \Pi\sigma\|_{L^2(\Omega)} + C \inf_{w^h \in \mathcal{W}^h} \|\varepsilon(u - w^h)\|_{L^2(\Omega)}.$$

Of course, the divergence error is just $e_2 = f - Pf$.

3.2. The projector.

The natural interpolant of σ into our finite element, denoted by $\Pi_K\sigma$ is defined by $\ell(\Pi_K\sigma) = \ell(\sigma)$ for all the degrees of freedom $\ell \in \{\ell_\rho, \ell_\mu, \ell_s\}$ in (2.5) or (2.5′). It is easy to see that this projector satisfies Assumption 3.2. The only difficulty is that the domain D_Π for this projector consists of functions with high smoothness requirements (to make the functionals ℓ_s continuous). Then requiring that the exact stress lies in D_Π (e.g., as in Theorem 3.3), is rather restrictive. Therefore we now define a different Π with an enlarged domain D_Π, closely following a construction in [4, 8].

We will first need to define an auxiliary projection following [4, 8] by zeroing out the edge degrees of freedom in three dimensions (and the vertex ones in the two dimensions). I.e., in the $d = 3$ case, define the function $\Pi^0\sigma|_K \in \mathcal{P}^{k+1}(K, \mathbb{S})$ for each K by

$$\ell_\rho(\Pi^0\sigma) = \ell_\rho(\sigma), \qquad \text{for all } \rho \in \mathcal{P}^{k-1}(K, \mathbb{S}), \tag{3.19a}$$

$$\ell_\mu(\Pi^0\sigma) = \ell_\mu(\sigma), \qquad \text{for all } \mu \in \mathcal{P}^k(F, \mathbb{R}^d), \text{ for all faces } F \text{ of } K, \tag{3.19b}$$

$$\ell_s(\Pi^0\sigma) = 0 \qquad \text{for all } s \in \mathcal{P}^{k+1}(e, \mathbb{R}), \text{ for all edges } e \text{ of } K. \tag{3.19c}$$

In the $d = 2$ case the definition is similar except that (3.19c) is replaced by

$$(\Pi^0\sigma(x))n^- \cdot n^+ = 0 \qquad \text{for all vertices } x \text{ of } K, \tag{3.19c′}$$

and F in (3.19b) denotes edges of the triangular element.

LEMMA 3.5: *The projection Π^0 satisfies Assumption 3.2 with $D_\Pi = H(\mathrm{div}, \Omega, \mathbb{S}) \cap L^p(\Omega; \mathbb{S})$ for any $p > 2$. Moreover, if σ is in $H^1(K, \mathbb{S})$, then*

$$\|\Pi^0\sigma\|_{L^2(K)} \leq C(\|\sigma\|_{L^2(K)} + h_K|\sigma|_{H^1(K)}), \tag{3.20}$$

for all $K \in \Omega_h$. Here C only depends on the shape regularity of Ω_h.

Proof. By Theorem 2.1 we know that $\Pi^0\sigma$ is well-defined for smooth σ. The fact that Π^0 is continuous on $H(\text{div},\Omega,\mathbb{S}) \cap L^p(\Omega;\mathbb{S})$ for any $p > 2$ follows from the fact that we only use interior and face degrees of freedom (see e.g., [12]).

To prove the commutativity (3.15), let $w \in \mathcal{P}^k(K,\mathbb{R}^d)$. Then, integrating by parts,

$$\begin{aligned}
(\text{div}_h \Pi^0\sigma, w)_K &= - (\Pi^0\sigma, \varepsilon(w))_K + \langle(\Pi^0\sigma)n, w\rangle_{\partial K} \\
&= - (\sigma, \varepsilon(w))_K + \langle\sigma n, w\rangle_{\partial K} \\
&= (\text{div}_h \sigma, w)_K = (P\text{div}_h \sigma, w)_K.
\end{aligned}$$

In the second equality we used (3.19b) with $\mu = w$ and (3.19a) with $\rho = \varepsilon(w)$.

It only remains to prove (3.20). We only consider the $d = 3$ case as the other is similar. Because of Theorem 2.1, it is easy to see by a scaling argument, using the same mappings as in [17, § 2.2], that

$$C\|\tau\|_{L^2(K)} \leq \sup_{\rho\in\mathcal{P}^{k-1}(K,\mathbb{S})} \frac{\ell_\rho(\tau)}{\|\rho\|_{L^2(K)}} + \sup_{F,\mu\in\mathcal{P}^k(F,\mathbb{R}^d)} \frac{h_K^{1/2}\ell_\mu(\tau)}{\|\mu\|_{L^2(F)}} + \sup_{e,\,s\in\mathcal{P}^{k+1}(e,\mathbb{R})} \frac{h_K\ell_s(\tau)}{\|s\|_{L^2(e)}}$$

holds for all τ in $\mathcal{P}^{k+1}(K,\mathbb{S})$. We apply this with $\tau = \Pi^0\sigma$. The last term then vanishes due (3.19c). The first term on the right hand side is bounded by $\|\sigma\|_{L^2(\Omega)}$ due to (3.19a). Using a trace inequality for the remaining term, $\ell_\mu(\Pi^0\sigma) = \ell_\mu(\sigma) \leq \|\mu\|_F(\|\sigma\|_{L^2(K)} + h_K|\sigma|_{H^1(K)})$. This proves the lemma. $\square$

The operator Π^0, although continuous on $H(\text{div},\Omega,\mathbb{S}) \cap L^p(\Omega;\mathbb{S})$ for $p > 2$, does not have good approximation properties. Therefore, following [8] we modify it further. Consider the Clement interpolant [14, 18] of order k. Let R^k denote its matrix version in which the Clement interpolant acts component wise. For any $0 \leq r \leq k$, we have the following local approximation result for the Clement interpolant (see [18])

$$\|R^k\sigma - \sigma\|_{L^2(K)} + h_K|R^k\sigma - \sigma|_{H^1(K)} \leq C\,h_K^{r+1}\|\sigma\|_{H^{r+1}(\triangle_K)} \tag{3.21}$$

where $\triangle_K$ is the union of simplices that share a vertex with K. The global interpolant is then defined by

$$\Pi = \Pi^0(I - R^{k+1}) + R^{k+1}. \tag{3.22}$$

THEOREM 3.6: *The projector Π defined above satisfies Assumption 3.2 with $H(\text{div},\Omega,\mathbb{S})\cap L^p(\Omega,\mathbb{S}) \subset D_\Pi$ for any $p > 2$. Moreover, for $0 \leq r \leq k+1$ we have*

$$\|\Pi\sigma - \sigma\|_{L^2(K)} \leq Ch_K^{r+1}\|\sigma\|_{H^{r+1}(\triangle_K)} \tag{3.23}$$

for all $K \in \Omega_h$ where C only depends on the shape regularity of Ω_h.

Proof. To prove the commutativity property (3.15) of Assumption 3.2, we first integrate by parts on each $K \in \Omega_h$ and observe that

$$(\text{div}_h(\Pi\sigma-\Pi^0\sigma), w)_K = -(R^{k+1}\sigma-\Pi^0 R^{k+1}\sigma, \varepsilon(w))_K + \langle(R^{k+1}\sigma - \Pi^0 R^{k+1}\sigma)n, w\rangle_{\partial K}$$

Using (3.19), this implies

$$(\text{div}_h(\Pi\sigma - \Pi^0\sigma), w)_K = 0, \quad \text{for all } w \in \mathcal{P}^k(K,\mathbb{R}^d).$$

Hence, $\text{div}_h \Pi\sigma = \text{div}_h \Pi^0\sigma = P\text{div}_h \sigma$ where we used Lemma 3.5. This proves (3.15).

In order to prove (3.23),

$$\|\Pi\sigma - \sigma\|_{L^2(K)} = \|(\Pi^0 - I)(R^{k+1} - I)\sigma\|_{L^2(K)}$$
$$\leq C(\|R^{k+1}\sigma - \sigma\|_{L^2(K)} + h_K|R^{k+1}\sigma - \sigma|_{H^1(K)}).$$

where we used (3.20) of Lemma 3.5. The result now follows from (3.21).   □

3.3. **Displacement Error.** In this section we prove an error estimate for $\|Pu - u^h\|_{L^2(\Omega)}$.

THEOREM 3.7: *If Π is given by (3.22) for $k \geq 1$, then whenever $\sigma \in D_\Pi$,*

$$C\|Pu - u^h\|_{L^2(\Omega)} \leq \|\sigma - \Pi\sigma\|_{L^2(\Omega)} + \inf_{w^h \in \mathcal{W}^h}(\|u - w^h\|_{L^2(\Omega)} + \|\varepsilon(u - w^h)\|_{L^2(\Omega)}).$$
(3.24)

Proof. There exists $\psi \in H^1(\Omega, \mathbb{S})$ satifying

$$\operatorname{div}\psi = Pu - u^h \quad \text{in } \Omega \tag{3.25}$$

with

$$\|\psi\|_{H^1(\Omega)} \leq C\|Pu - u^h\|_{L^2(\Omega)}. \tag{3.26}$$

$$\begin{aligned}
\|Pu - u^h\|_{L^2(\Omega)}^2 &= (Pu - u^h, \operatorname{div}_h\psi)_\Omega & \text{by (3.25)} \\
&= (Pu - u^h, \operatorname{div}_h\Pi\psi)_\Omega & \text{by (3.15)} \\
&= (u - u^h, \operatorname{div}_h\Pi\psi)_\Omega & \text{by definition of } P \\
&= -(\mathcal{A}(\sigma - \sigma^h), \Pi\psi)_\Omega + \langle u, \Pi\psi n\rangle_{\partial\Omega_h} & \text{by (3.13a).}
\end{aligned}$$

The first term can easily be estimated as follows

$$\begin{aligned}
-(\mathcal{A}(\sigma - \sigma^h), \Pi\psi)_\Omega &\leq C\|\sigma - \sigma^h\|_{L^2(\Omega)}\|\Pi\psi\|_{L^2(\Omega)} \\
&\leq C\|\sigma - \sigma^h\|_{L^2(\Omega)}\|\psi\|_{H^1(\Omega)} \\
&\leq C\|\sigma - \sigma^h\|_{L^2(\Omega)}\|Pu - u^h\|_{L^2(\Omega)},
\end{aligned}$$

where we used (3.23) and (3.26).

To bound the second term we note that $\langle w^h, \Pi\psi n\rangle_{\partial\Omega_h} = 0$ for any $w^h \in \mathcal{W}^h$ since $\Pi\psi \in V^h$. Hence,

$$\begin{aligned}
\langle u, \Pi\psi n\rangle_{\partial\Omega_h} &= \langle u - w^h, \Pi\psi n\rangle_{\partial\Omega_h} \\
&= (\varepsilon(u - w^h), \Pi\psi)_\Omega + (u - w^h, \operatorname{div}_h(\Pi\psi))_\Omega \\
&= (\varepsilon(u - w^h), \Pi\psi)_\Omega + (u - w^h, Pu - u^h)_\Omega,
\end{aligned}$$

where we used integration by parts, (3.15) and (3.25).

Therefore, we see that

$$\langle u, \Pi\psi n\rangle_{\partial\Omega_h} \leq C(\|u - w^h\|_{L^2(\Omega)} + \|\varepsilon(u - w^h)\|_{L^2(\Omega)})\|Pu - u^h\|_{L^2(\Omega)},$$

where we used (3.23) and (3.26).

Hence, we obtain

$$C\|Pu - u^h\|_{L^2(\Omega)} \leq \|\sigma - \sigma^h\|_{L^2(\Omega)} + \inf_{w^h \in \mathcal{W}^h}(\|u - w^h\|_{L^2(\Omega)} + \|\varepsilon(u - w^h)\|_{L^2(\Omega)}).$$
(3.27)

To complete the proof we use Theorem 3.3. □

3.4. **Duality argument.** In this section we prove an inequality that leads to the second part, namely (3.12), of Theorem 3.1. But first, let us clarify what we meant by full regularity there. Consider the dual problem of finding ψ in $H(\text{div}, \Omega, \mathbb{S})$ and ϕ in $H^1(\Omega, \mathbb{R}^d)$ satisfying

$$\text{div}\,\psi = \theta \quad \text{in } \Omega, \tag{3.28a}$$
$$\mathcal{A}\psi - \varepsilon(\phi) = 0 \quad \text{in } \Omega, \tag{3.28b}$$
$$\phi = 0 \quad \text{on } \partial\Omega, \tag{3.28c}$$

for any given θ in $L^2(\Omega, \mathbb{R}^d)$. We have "full elliptic regularity" if the solution of the dual problem satisfies

$$\|\psi\|_{H^1(\Omega)} + \|\phi\|_{H^2(\Omega)} \le C\|\theta\|_{L^2(\Omega)}. \tag{3.29}$$

for all $\theta \in L^2(\Omega, \mathbb{R}^d)$. This is known to hold in many instances, e.g., see [11] for convex polygons.

THEOREM 3.8: *Assuming full elliptic regularity (3.29) we have for any $k \ge 1$,*

$$C\|Pu - u^h\|_{L^2(\Omega)} \le h\|\sigma - \sigma^h\|_{L^2(\Omega)} + h \inf_{w^h \in \mathcal{W}^h} \|\varepsilon(u - w^h)\|_{L^2(\Omega)}.$$

Proof. Let $\theta = Pu - u^h$ in (3.28). Then,

$$\begin{aligned}
\|Pu - u^h\|_{L^2(\Omega)}^2 &= (Pu - u^h, \text{div}_h\,\psi)_\Omega & \text{by (3.28a)}\\
&= (Pu - u^h, \text{div}_h\,\Pi\psi)_\Omega & \text{by (3.15)}\\
&= (u - u^h, \text{div}_h\,\Pi\psi)_\Omega & \text{by definition of } P\\
&= -(\mathcal{A}(\sigma - \sigma^h), \Pi\psi)_\Omega + \langle u, \Pi\psi n\rangle_{\partial\Omega_h} & \text{by (3.13a)}\\
&= -(\mathcal{A}(\sigma - \sigma^h), \Pi\psi - \psi)_\Omega + \langle u, \Pi\psi n\rangle_{\partial\Omega_h}\\
&\quad -(\mathcal{A}(\sigma - \sigma^h), \psi)_\Omega.
\end{aligned}$$

We rewrite the last term as follows.

$$\begin{aligned}
(\mathcal{A}(\sigma - \sigma^h), \psi)_\Omega &= (\sigma - \sigma^h, \varepsilon(\phi))_\Omega & \text{by (3.28b)}\\
&= (\sigma - \sigma^h, \varepsilon(\phi - \phi^h))_\Omega
\end{aligned}$$

for any ϕ^h in $\mathcal{W}^h$. The last equality holds because by integration by parts

$$(\sigma - \sigma^h, \varepsilon(\phi^h))_\Omega = \langle(\sigma - \sigma^h)n, \phi^h\rangle_{\partial\Omega_h} - (\text{div}_h\,(\sigma - \sigma^h), \phi^h)_\Omega,$$

and both terms on the right hand side are zero. We see that the first is zero by the continuity and weak continuity of σn and $\sigma^h n$, resp., across mesh faces (and noting that ϕ^h is single-valued there). The second vanishes by (3.13b). Hence,

$$\|Pu - u^h\|_{L^2(\Omega)}^2 = -(\mathcal{A}(\sigma - \sigma^h), \Pi\psi - \psi)_\Omega + \langle u, \Pi\psi n\rangle_{\partial\Omega_h} - (\sigma - \sigma^h, \varepsilon(\phi - \phi^h))_\Omega$$

Let us name the terms on the right hand side consecutively as T_1, T_2, and T_3.

We now estimate these three terms individually. Using the Cauchy-Schwarz inequality, (3.23), and (3.29) we obtain

$$T_1 \le C\,h\,\|\sigma - \sigma^h\|_{L^2(\Omega)}\|\psi\|_{H^1(\Omega)} \le C\,h\,\|\sigma - \sigma^h\|_{L^2(\Omega)}\,\|Pu - u^h\|_{L^2(\Omega)}.$$

For T_2, we note that for any w^h in $\mathcal{W}^h$ $\langle w^h, \Pi\psi n\rangle_{\partial\Omega_h} = 0$ since $\Pi\psi \in V^h$. Hence,

$$T_2 = \langle u - w^h, \Pi\psi n\rangle_{\partial\Omega_h}.$$

Furthermore,

$$T_2 = \langle u - w^h, \Pi\psi n - \psi n\rangle_{\partial\Omega_h},$$

which follows since both $u - w^h$ and ψn are single-valued on mesh faces and $u - w^h$ vanishes on $\partial\Omega$. Integrating by parts and using (3.15),

$$T_2 = (\varepsilon(u - w^h), \Pi\psi - \psi)_\Omega \le \|\varepsilon(u - w^h)\|_{L^2(\Omega)}\|\Pi\psi - \psi\|_{L^2(\Omega)}.$$

Here we have used that $\operatorname{div}_h(\Pi\psi - \psi) = P\theta - \theta = 0$ since $\theta \in W_h$. Combining (3.23) with (3.29) we get

$$T_2 \le C\, h\|\varepsilon(u - w^h)\|_{L^2(\Omega)}\|Pu - u^h\|_{L^2(\Omega)}.$$

For T_3, we select a ϕ^h in $\mathcal{W}^h$ with good approximation properties, e.g., the one provided by the Scott-Zhang interpolation operator [27]. Let $Z^k : H^1(\Omega, \mathbb{R}^d) \to \mathcal{W}^h$ denote this interpolation of degree k, applied component by component to vector functions. Then by the results in [27],

$$\|\phi - Z^k\phi\|_{L^2(K)}^2 + h_K\|\operatorname{grad}(\phi - Z^k\phi)\|_{L^2(K)}^2 \le C\, h^{r+1}\|\phi\|_{H^{r+1}(\triangle_K)}, \qquad (3.30)$$

where $\triangle_K$ is the union of simplices that share a vertex with K. Set $\phi^h = Z^k\phi$ in T_3. Using (3.30) and (3.29), we then have

$$T_3 \le C\, h\, \|\sigma - \sigma^h\|_{L^2(\Omega)}\|Pu - u^h\|_{L^2(\Omega)}.$$

Combining the bounds for T_1, T_2, T_3, the proof is finished. $\square$

We conclude with the proof of the previously stated main result of this section.

Proof of Theorem 3.1. To prove the first estimate (3.11), we apply Theorem 3.3 and Theorem 3.7 and use the triangle inequality. The assumption in Theorem 3.3 on the projector has been verified by Theorem 3.6. The infimum in (3.27) and (3.24) can be bounded by setting $w^h = Z^k u$ and using (3.30).

The second estimate (3.12) of the theorem follows from Theorem 3.8, the triangle inequality and using the same choice of w^h. $\square$

4. Hybrid Form

We give an alternative formulation of the method (1.3) which results in a symmetric positive definite system for a single new variable. All the original variables can be locally recovered after solving for this new variable. The alternative formulation is obtained by hybridization, which removes the interelement continuity requirements from the space V^h, and places them as an additional equation of the method. Accordingly, we need the space

$$\tilde{V}^h = \{\tau : \tau|_K \in \mathcal{P}^{k+1}(K, \mathbb{S}) \text{ for all mesh elements } K \in \Omega_h\}$$

without any interelement continuity constraints, as well as a space of Lagrange multipliers

$$\Lambda^h = \{\mu : \mu|_F \in \mathcal{P}^k(F, \mathbb{R}^d) \text{ for all mesh faces } F \text{ of } \Omega_h, \text{ and } \mu|_{\partial\Omega} = 0\}.$$

The approximate solution given by the hybridized method is $(\sigma^h, u^h, \lambda^h) \in \tilde{V}^h \times W^h \times \Lambda^h$, satisfying

$$(\mathcal{A}\sigma^h, \tau)_\Omega + (u^h, \mathrm{div}_h\, \tau)_\Omega + \langle \lambda^h, \tau n \rangle_{\partial\Omega_h} = 0, \tag{4.31a}$$

$$(\mathrm{div}_h\, \sigma^h, \omega)_\Omega = (f, \omega)_\Omega, \tag{4.31b}$$

$$\langle \sigma^h n, \mu \rangle_{\partial\Omega_h} = 0, \tag{4.31c}$$

for all $(\tau, \omega, \mu) \in \tilde{V}^h \times W^h \times \Lambda^h$. Here, the notation $\langle \cdot, \cdot \rangle_{\partial\Omega_h}$ is as in (**??**).

PROPOSITION 4.1: *There is a unique $(\sigma^h, u^h, \lambda^h) \in \tilde{V}^h \times W^h \times \Lambda^h$ satisfying (4.31). Moreover, the first two components of the solution coincide with that of the mixed method (1.3).*

Proof. If $(\sigma^h, u^h, \lambda^h)$ satisfies (4.31), then (4.31c) give us that $\sigma^h \in V^h$. Moreover, since $V^h \subset \tilde{V}^h$, choosing test functions $\tau \in V^h$, we see that the equations (4.31a)–(4.31b) are identical to the equations of the mixed method (1.3). Therefore, (σ^h, u^h) solves (1.3).

Next we prove that (4.31) has a unique solution. Since the system (4.31) is square, it is enough to prove uniqueness. If f is identically zero, the argument of the previous paragraph and Corollary 2.4 shows that (σ^h, u^h) vanishes. The equation (4.31a) then becomes

$$\langle \lambda^h, \tau n \rangle_{\partial\Omega_h} = 0 \qquad \text{for all } \tau \in \tilde{V}^h.$$

We then use the face degrees of freedom in (2.5b) (or the edge degrees in the two-dimensional case (2.5b′)) to show that λ^h also vanishes. $\square$

We next show that it is possible to eliminate the variables σ^h and u^h, and obtain a global linear system solely for λ^h. Let $A : \tilde{V}^h \mapsto \tilde{V}^h$ and $B : \tilde{V}^h \mapsto W^h$ and $C : \tilde{V}^h \mapsto \Lambda^h$ be defined by

$$(A\sigma, \tau)_{\Omega_h} = (\mathcal{A}\sigma, \tau)_{\Omega_h}, \qquad (B\sigma, w)_{\Omega_h} = (w, \mathrm{div}_h\, \sigma)_{\Omega_h}, \qquad \langle C\sigma, \mu \rangle_{\partial\Omega_h} = \langle \mu, \sigma n \rangle_{\partial\Omega_h},$$

for all $\sigma, \tau \in \tilde{V}^h$, $w \in W^h$, and $\mu \in \Lambda^h$. Denoting their adjoints by superscript ′, the hybridized system (4.31) can be rewritten as

$$\begin{pmatrix} A & B' & C' \\ B & 0 & 0 \\ C & 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma^h \\ u^h \\ \lambda^h \end{pmatrix} = \begin{pmatrix} 0 \\ F^h \\ 0 \end{pmatrix} \tag{4.32}$$

where F^h is the L^2-orthogonal projection of f into W^h. Such systems were considered for hybridization abstractly in [16, Appendix A] (see also [15]) under the assumption that B is surjective. To see that this assumption holds for our B, given any w^h in W^h, we put $f = w^h$ and solve the mixed method (1.3). There is a unique solution by Corollary 2.4. The resulting σ^h satisfies $B\sigma^h = w^h$, so B is surjective.

A consequence of this surjectivity is that there is a unique $(\sigma_{F,G}, u_{F,G}) \in \tilde{V}^h \times W^h$ solving

$$\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \begin{pmatrix} \sigma_{F,G} \\ u_{F,G} \end{pmatrix} = \begin{pmatrix} G \\ F \end{pmatrix}. \tag{4.33}$$

Note that $(\sigma_{F,G}, u_{F,G})$ can be computed locally, element by element, as $\tilde{V}^h$ has no interelement continuity. Let (σ_μ, u_μ) denote the $(\sigma_{F,G}, u_{F,G})$ obtained by setting $G =$

$-C'\mu$ and $F = 0$, and let (σ_f, u_f) denote the $(\sigma_{F,G}, u_{F,G})$ obtained when $G = 0$ and $F = F^h$.

THEOREM 4.2: *The function* $(\sigma^h, u^h, \lambda^h) \in \tilde{V}^h \times W^h \times \Lambda^h$ *satisfies* (4.31) *if and only if* λ^h *is the unique solution of*

$$a_h(\lambda^h, \mu) = b_h(\mu), \qquad \text{for all } \mu \in \Lambda^h, \tag{4.34}$$

where the forms are defined by $a_h(\mu, \gamma) := (\mathcal{A}\, \sigma_\mu, \sigma_\gamma)_{\Omega_h}$, $b_h(\mu) := (f, u_\mu)_{\Omega_h}$. *Furthermore,*

$$\sigma^h = \sigma_{\lambda^h} + \sigma_f, \tag{4.35a}$$

$$u^h = u_{\lambda^h} + u_f. \tag{4.35b}$$

Proof. Since B is surjective, we can apply [16, Theorem A.1]. $\qquad\square$

Equation (4.34) gives a symmetric positive definite system for λ^h. In practical implementations, this may be preferable over a direct assembly of (1.3). The latter will result in a larger indefinite system. Moreover, due to the (already mentioned) local nature of (4.33), the right hand sides of (4.35) are locally computable from λ^h and f, so once the (global) positive definite system for λ^h is solved, the approximate stress and displacement can be recovered locally.

5. A REDUCED ELEMENT

This section provides an answer to the following natural question: Is it possible to reduce the stress space $\mathcal{P}^{k+1}(K, \mathbb{S})$ and yet maintain the same order of convergence?

We first consider the three dimensional case. For any edge e of a tetrahedron K, let e^* denote the "opposite" edge, i.e., none of the two faces that share e^* have e as one of their edges. Let $F(e^*)$ denote *any one* of the two faces that share e^*. Define the reduced space of stresses by

$$V(K) = \left\{ \sigma \equiv \sum_e p_e\, t_e t'_e \ \middle|\ p_e \in \mathcal{P}_e \right\} \tag{5.36}$$

where the sum runs over all six edges e of K, t_e denotes a tangent vector along an edge e, and $\mathcal{P}_e = \{p \in \mathcal{P}^{k+1}(K, \mathbb{R}) : p|_{F(e^*)} \in \mathcal{P}^k(F(e^*), \mathbb{R})\}$. A unisolvent set of degrees of freedom for this space is furnished by (2.5) after we omit the edge degrees of freedom there, as the next theorem states.

THEOREM 5.1: *The set of linear functionals consisting of*

ℓ_ρ *defined in* (2.5a), $\qquad$ *for all* $\rho \in \mathcal{P}^{k-1}(K, \mathbb{S})$,

ℓ_μ *defined in* (2.5b), $\qquad$ *for all* $\mu \in \mathcal{P}^k(F, \mathbb{R}^d)$, *for all faces* F *of* K,

form a set of unisolvent degrees of freedom for $V(K)$.

We prove this theorem using two lemmas. Let $\mathcal{P}^k_\perp = \{p \in \mathcal{P}^k(K, \mathbb{R}) : (p, q_{k-1})_K = 0$ for all q_{k-1} in $\mathcal{P}^{k-1}(K, \mathbb{R})\}$ and $Q^{k+1}_{e,\perp} = \{\lambda_{e^*} r_k : r_k \in \mathcal{P}^k_\perp\}$, where λ_{e^*} denotes the barycentric coordinate which vanishes on $F(e^*)$.

LEMMA 5.2: $\mathcal{P}_e = \mathcal{P}^k(K, \mathbb{R}) \oplus Q^{k+1}_{e,\perp}$.

Proof. Clearly, $\mathcal{P}^k(K,\mathbb{R}) \oplus \mathcal{Q}^{k+1}_{e,\perp} \subseteq \mathcal{P}_e$, so we only need to prove the reverse inclusion. Since, any $p \in \mathcal{P}_e$ when restricted to $F(e^*)$, is of degree at most k, there exists a q_k in $\mathcal{P}^k(K,\mathbb{R})$ such that

$$p - q_k = 0, \qquad \text{on } F(e^*).$$

Thus, $p - q_k = \lambda_{e^*} w_k$ for some w_k in $\mathcal{P}^k(K,\mathbb{R})$. Decomposing $w_k = w_{k-1} + w_\perp$ with $w_{k-1} \in \mathcal{P}^{k-1}(K,\mathbb{R})$ and $w_\perp \in \mathcal{P}^k_\perp$, we find that

$$p = \tilde{q}_k + \lambda_{e^*} w_\perp,$$

with $\tilde{q}_k = q_k + \lambda_{e^*} w_{k-1} \in \mathcal{P}^k(K,\mathbb{R})$. Thus $\mathcal{P}_e = \mathcal{P}^k(K,\mathbb{R}) + \mathcal{Q}^{k+1}_{e,\perp}$.

The decomposition is direct, because if $\lambda_{e^*} q_\perp = p_k$ for some $q_\perp$ in $\mathcal{P}^k_\perp$ and $p_k \in \mathcal{P}^k(K,\mathbb{R})$, then p_k vanishes on $F(e^*)$, so $p_k = \lambda_{e^*} r_{k-1}$ for some r_{k-1} in $\mathcal{P}^{k-1}(K,\mathbb{R})$. Thus,

$$(q_\perp, q_\perp)_K = (\lambda^{-1}_{e^*} \lambda_{e^*} q_\perp, q_\perp)_K = (\lambda^{-1}_{e^*} p_k, q_\perp)_K = (r_{k-1}, q_\perp)_K = 0,$$

so $q_\perp$ and p_k vanish. $\qquad\square$

The next observation is a simple identity. We use the same notations introduced in the proof of Theorem 2.1 such as λ_i, n_i, F_i, etc., for $i = 1, 2, 3, 4$. Additionally, now we let e_{ij} denote the edge connecting the vertices where λ_i and λ_j equal one. We assume that the index set $\{i, j, l, m\}$ is a permutation of $\{1, 2, 3, 4\}$.

LEMMA 5.3: *For any $\sigma \in V(K)$ as in (5.36), and $i \neq j$, the following identity holds on all points in K.*

$$\sigma n_i \cdot n_j = p_{e_{ij}}(t_{e_{ij}} \cdot n_i)(t_{e_{ij}} \cdot n_j)$$

Proof. Clearly, n_i is orthogonal to the three tangent vectors $t_{e_{lm}}, t_{e_{mj}}, t_{e_{jl}}$ on F_i. Similarly n_j is orthogonal to $t_{e_{lm}}, t_{e_{mi}}, t_{e_{il}}$. Hence in the sum

$$\sigma n_i \cdot n_j = \sum_e p_e (t_e \cdot n_i)(t_e \cdot n_j)$$

only one summand is nonzero, and this is precisely the term stated in the lemma. $\quad\square$

Proof of Theorem 5.1. First we must count $\dim(V(K))$. To this end we first note that (5.36) can rewritten as

$$V(K) = \underset{e}{\oplus} (t_e t'_e) \mathcal{P}_e \tag{5.37}$$

To see that the above sum is direct, suppose $\sigma = \sum_e p_e t_e t'_e \equiv 0$. Then using Lemma 5.3 with all combinations of distinct i and j, we find that $p_e \equiv 0$ for all edges e. Hence

$$\dim(V(K)) = 6 \dim(\mathcal{P}_e).$$

By Lemma 5.2, $\dim(\mathcal{P}_e) = \dim(\mathcal{P}^k(K,\mathbb{R})) + \dim(\mathcal{P}^k_\perp)$. Simplifying, we obtain

$$\dim(V(K)) = (k+2)(k+3)(k+4) - 6(k+2),$$

which matches the number of degrees of freedom given in the theorem.

Next, suppose $\sigma \in V(K)$ satisfies $\ell_\mu(\sigma) = \ell_\rho(\sigma) = 0$. We will show that σ vanishes. Pick any two faces F_i and F_j. They share the edge e_{lm}. Obviously, $e_{lm} = e^*_{ij}$. Hence the restriction of $p_{e_{ij}}$ to one of the faces F_i or F_j is of degree at most k. Without

loss of generality, let that face be F_i. Then, $\mu = n_j p_{e_{ij}}|_{F_i} \in \mathcal{P}^k(F_i, \mathbb{R}^3)$. Applying Lemma 5.3,

$$0 = \langle \sigma n_i, \mu \rangle_{F_i} = (t_{e_{ij}} \cdot n_i)(t_{e_{ij}} \cdot n_j) \int_{F_i} p_{e_{ij}}^2.$$

Hence $p_{e_{ij}}$ vanishes on F_i. Consequently, there exists a $q_k \in \mathcal{P}^k(K, \mathbb{R})$ such that $p_{e_{ij}} = \lambda_i q_k$. Now, considering the remaining face F_j, setting $\mu = n_i q_k|_{F_j}$, and applying Lemma 5.3, we find that

$$0 = \langle \sigma n_j, \mu \rangle_{F_j} = (t_{e_{ij}} \cdot n_i)(t_{e_{ij}} \cdot n_j) \int_{F_j} \lambda_i q_k^2.$$

Hence q_k must vanish on F_j and consequently there exists a w in $\mathcal{P}^{k-1}(K, \mathbb{R})$ such that

$$p_{e_{ij}} = \lambda_i \lambda_j w.$$

Now, using the interior degrees of freedom (cf. (2.6)) we find that $p_{e_{ij}} \equiv 0$. Since i and j were two arbitrary distinct indices, σ vanishes. $\qquad \square$

We would like to note that using Lemma 5.2 the space $V(K)$ given in (5.36) can be written as

$$V(K) = \mathcal{P}^k(K, \mathbb{S}) + \left\{ \sigma \equiv \sum_e q_e\, t_e t'_e \;\middle|\; q_e \in \mathcal{Q}_{e,\perp}^{k+1} \right\}. \tag{5.38}$$

The two dimensional case is similar to the three dimensional case. For each edge e of the triangle K we let t_e denote the tangent vector to e. We let e^* be *any one* of the two edges of K that is not equal to e. Similarly, $\mathcal{Q}_{e,\perp}^{k+1} = \{\lambda_{e^*} r_k : r_k \in \mathcal{P}_\perp^k\}$, where λ_{e^*} denotes the barycentric coordinate which vanishes on e^*. We define $V(K)$ in two dimensions by (5.38) where the sum is taken over the three edges of the triangle K. The degrees of freedom (2.5a') and (2.5b') form a unisolvent set of degrees of freedom for the space $V(K)$. The proof is similar to the three dimensional case, so we leave the details to the reader.

Finally, consider the mixed method (1.3) with the new stress space, i.e., now V^h is as in (1.2a), except $\mathcal{P}^{k+1}(K, \mathbb{S})$ is now replaced with $V(K)$. The displacement space W^h remains the same. This method can be analyzed as in § 3. Since $\mathcal{P}^k(K, \mathbb{S}) \subseteq V(K)$, the stress space has the required approximation properties for the analysis. So as not to repeat the details of previous sections, we simply summarize the results:

(1) The statements of Theorem 3.1 hold for the solution of the mixed noncon-
    forming method using the above reduced space.
(2) The new method can also be hybridized as in § 4.

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.

[2] M. Amara and J. M. Thomas, *Equilibrium finite elements for the linear elastic problem*, Numer. Math., 33 (1979), pp. 367–383.

[3] D.N. Arnold, G. Awanou *Rectangular mixed finite elements for elasticity*, Math. Models Methods Appl. Sci. 15 (2005), no. 9, 1417–1429.

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

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

[6] D. Arnold, J. Douglas, C. Gupta, *A family of higher order mixed finite element methods for plane elasticity*, Numer. Math. 45 (1984), no. 1, 1–22.

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

[8] D. N. Arnold and R. Winther , *Mixed finite elements for elasticity*, Numer. Math., 92 (2002), pp. 401–419.

[9] D. N. Arnold and R. Winther, *Nonconforming mixed elements for elasticity*, Dedicated to Jim Douglas, Jr. on the occasion of his 75th birthday. Math. Models Methods Appl. Sci. 13 (2003), no. 3, 295–307.

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

[11] 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), pp. 874–878.

[12] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

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

[14] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO (Revue Française d'Atomatique, Informatique et Recherche Opérationnelle), Analyse Numérique, R-2 (9e année) (1975), pp. 77–84.

[15] B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for the Dirichlet problem*, SIAM J. Numer. Anal., 42 (2004), pp. 283–301.

[16] B. Cockburn and J. Gopalakrishnan, *Error analysis of variable degree mixed methods for elliptic problems via hybridization*, Math. Comp., 74 (2005), pp. 1653–1677 (electronic).

[17] B. Cockburn, J. Gopalakrishnan, and J. Guzmán, *A new elasticity element made for enforcing weak stress symmetry*, Math. Comp., 79 (2010), pp. 1331–1349.

[18] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, Applied Mathematical Sciences, 159. Springer-Verlag, New York, 2004.

[19] M. Farhloul and M. Fortin, *Dual hybrid methods for the elasticity and the Stokes problems: a unified approach*, Numer. Math., 76 (1997), pp. 419–440.

[20] J. Gopalakrishnan, and J. Guzmán, *A second elasticity method with tightened stress symmetry*, To appear in IMA J. Numer. Anal., (2010).

[21] J. Guzmán, *A unified analysis of several mixed methods for elasiticity with weak stress symmetry*, submitted.

[22] J. Hu and Z.-C. Shi, *Lower order rectangular nonconforming mixed finite elements for plane elasticity*, SIAM J. Numer. Anal. 46 (2007/08), no. 1, 88–102.

[23] 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.

[24] C. Johnson and B. Mercier, *Some equilibrium finite element methods for two-dimensional elasticity problems*, Numer. Math. 30 (1978), no. 1, 103–116.

[25] M. Morley, *A family of mixed finite elements for linear elasticity* Numer. Math. 55 (1989), no. 6, 633–666.

[26] W. Qiu and L. Demkowicz, *Mixed hp-finite element method for linear elasticity with weakly imposed symmetry*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 3682–3701.

[27] R.L. Scott, S. Zhang, *Finite element interpolation of non-smooth functions satisfying boundary conditions*, Math. Comp. 54 (1990), no. 190, 483–493.

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

[29] 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.

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

Department of Mathematics, University of Florida, Gainesville, FL 32611–8105
E-mail address: jayg@ufl.edu

Division of Applied Mathematics, Brown University, Providence, RI 02912
E-mail address: johnny_guzman@brown.edu