

# A SECOND ELASTICITY ELEMENT USING THE MATRIX BUBBLE

J. GOPALAKRISHNAN AND J. GUZMÁN

ABSTRACT. We presented a family of finite elements that use a polynomial space augmented by certain matrix bubbles in [Math. Comp., 79 (2010), 1331–1349]. In this sequel, we exhibit a second family of elements that use the same matrix bubble. This second element uses a stress space smaller than the first, while maintaining the same space for rotations (which are the Lagrange multipliers corresponding to a weak symmetry constraint). The space of displacements are of one degree less than the first method. The analysis, while similar to the first, requires a few new adjustments as the new Fortin projector may not preserve weak symmetry, but we are able to prove optimal convergence for all the variables. Finally, we present a sufficient condition wherein a mixed method with weakly imposed stress symmetry in fact yields an exactly symmetric stress tensor approximation.

## 1. INTRODUCTION

We introduce a close relative of the family of stress finite elements presented in [15]. These elements are applied to numerically solve the linear elasticity boundary value problem in a mixed formulation. While the mixed method is attractive in that it gives a direct approximation to the stress tensor  $\underline{\sigma}$ , generating symmetric stress tensor approximations poses significant difficulties. The symmetry is a requirement arising from the conservation of angular momentum, so it is of some considerable interest to preserve it at the discrete level. While it may be preferable to use finite elements that keep the exact symmetry of the stress tensor, the known symmetric stress elements for the mixed formulation have too many degrees of freedom to make it practically attractive [1, 3, 7]. Hence researchers have pursued avenues for computing weakly symmetric stress tensor approximations, by constraining various selected moments of  $\underline{\sigma} - \underline{\sigma}^t$  to vanish [5, 12, 15, 24, 19]. Our present contribution deals with a mixed formulation appropriate for weakly imposing symmetry of stress tensors. However, we also discuss a case where weak symmetry happens to imply exact symmetry.

The new stress finite element space we consider in this paper, in the three-dimensional (3D) case, is given by

$$\underline{\mathbf{V}}(K) = \underline{\mathcal{P}}^k(K) + \text{curl}((\text{curl} \tilde{\underline{\mathbf{A}}}^k(K)) \underline{\mathbf{b}}_K) \quad (1.1)$$

where  $K$  is a tetrahedron,  $\mathcal{P}^k(K)$  is the space of polynomials of degree  $k$  on  $K$ ,  $\underline{\mathcal{P}}^k(K) = [\mathcal{P}^k(K)]^{d \times d}$ , and  $\underline{\mathbf{b}}_K$  is “matrix bubble” introduced in [15], defined by

$$\underline{\mathbf{b}}_K := \sum_{\ell=0}^3 \lambda_{\ell-3} \lambda_{\ell-2} \lambda_{\ell-1} (\text{grad } \lambda_\ell)^t \text{grad } \lambda_\ell. \quad (1.2)$$

---

2000 *Mathematics Subject Classification.* 65M60,65N30,35L65.

*Key words and phrases.* finite element, elasticity, weakly imposed symmetry, mixed method.

This work was supported by the National Science Foundation under grants DMS-1014817 and DMS-0914596.

Here  $\lambda_j$  are the barycentric coordinates of  $K$  and the indices are calculated mod 4. Note that differential operators like curl and grad above are applied row-wise, so in particular  $\text{grad } \lambda_\ell$  is a *row* vector of derivatives of  $\lambda_\ell$ , so that  $(\text{grad } \lambda_\ell)^t \text{grad } \lambda_\ell$  is a  $3 \times 3$  matrix. In (1.1), we also used the space of anti-symmetric matrix polynomials  $\tilde{\mathbf{A}}^k(K)$  defined by

$$\begin{aligned}\underline{\mathbf{A}}^k(K) &= \{\underline{\boldsymbol{\eta}} \in \underline{\mathcal{P}}^k(K) : \underline{\boldsymbol{\eta}} + \underline{\boldsymbol{\eta}}^t = 0\}, \\ \tilde{\mathbf{A}}^k(K) &= \underline{\mathbf{A}}^k(K) \cap \underline{\mathcal{P}}_\perp^k(K),\end{aligned}$$

where  $\underline{\mathcal{P}}_\perp^k(K) := [\underline{\mathcal{P}}_\perp^k(K)]^{d \times d}$  and  $\underline{\mathcal{P}}_\perp^k(K) := \{v \in \underline{\mathcal{P}}^k(K) : (v, w)_K = 0 \text{ for all } w \in \underline{\mathcal{P}}^{k-1}(K)\}$ . Because of the properties of the matrix bubble (detailed later) the space added to  $\underline{\mathcal{P}}^k(K)$  in (1.1) is a space of ‘normal bubbles’ in the sense that they have zero normal stress components. This was one of the original motivations in the design of  $\underline{\mathbf{b}}_K$  (see [15] for details). Indeed, the results of [18] imply that multiplication of matrix polynomials with  $\underline{\mathbf{b}}_K$  gives matrices whose row-vectors are in a Nédélec subspace with zero tangential components, so their curls have zero normal components.

To compare with previously considered elements, let us introduce the mixed approximation to the elasticity problem with kinematic boundary conditions,

$$\text{div } \underline{\boldsymbol{\sigma}} = \underline{\mathbf{f}} \quad \text{in } \Omega, \quad (1.3a)$$

$$\mathcal{A} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\epsilon}}(\underline{\mathbf{u}}) = \mathbf{0} \quad \text{in } \Omega, \quad (1.3b)$$

$$\underline{\mathbf{u}} = \mathbf{0} \quad \text{on } \partial\Omega, \quad (1.3c)$$

where  $\mathcal{A}$  is a symmetric positive definite tensor over symmetric matrices and  $\underline{\mathbf{f}}$  is a given load function. The mixed approximations to the stress tensor ( $\underline{\boldsymbol{\sigma}}$ ), the displacement ( $\underline{\mathbf{u}}$ ), and the rotation ( $\underline{\boldsymbol{\rho}}$ ), form a triple  $(\underline{\boldsymbol{\sigma}}^h, \underline{\mathbf{u}}^h, \underline{\boldsymbol{\rho}}^h) \in \underline{\mathbf{V}}^h \times \underline{\mathbf{W}}^h \times \underline{\mathbf{A}}^h$  where the global finite element spaces are of the form

$$\underline{\mathbf{V}}^h := \{\underline{\mathbf{v}} \in \underline{\mathbf{H}}(\text{div}, \Omega) : \underline{\mathbf{v}}|_K \in \underline{\mathbf{V}}(K), \text{ for all } K \in \Omega_h\}, \quad (1.4a)$$

$$\underline{\mathbf{W}}^h := \{\underline{\mathbf{w}} \in \underline{\mathbf{L}}^2(\Omega) : \underline{\mathbf{w}}|_K \in \underline{\mathbf{W}}(K), \text{ for all } K \in \Omega_h\}, \quad (1.4b)$$

$$\underline{\mathbf{A}}^h := \{\underline{\mathbf{v}} \in \underline{\mathbf{L}}^2(\Omega) : \underline{\mathbf{v}}|_K \in \underline{\mathbf{A}}(K), \text{ for all } K \in \Omega_h\}. \quad (1.4c)$$

Here  $\Omega$  is a polygonal domain and  $\Omega_h$  is a simplicial finite element subdivision of  $\Omega$ .

Different authors have chosen different local finite element spaces  $\underline{\mathbf{V}}(K)$ ,  $\underline{\mathbf{W}}(K)$  and  $\underline{\mathbf{A}}(K)$ , as shown in Table 1, which also shows our choice in this paper. In all cases, the mixed approximation satisfies the system

$$(\mathcal{A} \underline{\boldsymbol{\sigma}}^h, \underline{\mathbf{v}})_\Omega + (\underline{\mathbf{u}}^h, \text{div } \underline{\mathbf{v}})_\Omega + (\underline{\boldsymbol{\rho}}^h, \underline{\mathbf{v}})_\Omega = 0 \quad (1.5a)$$

$$(\text{div } \underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\omega}})_\Omega = (\underline{\mathbf{f}}, \underline{\boldsymbol{\omega}})_\Omega \quad (1.5b)$$

$$(\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\eta}})_\Omega = 0, \quad (1.5c)$$

for all  $(\underline{\mathbf{v}}, \underline{\boldsymbol{\omega}}, \underline{\boldsymbol{\eta}}) \in \underline{\mathbf{V}}^h \times \underline{\mathbf{W}}^h \times \underline{\mathbf{A}}^h$ . Above and throughout, the notation  $(\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\theta}})_D$  denotes the integral over  $D$  of the dot product of the vector functions  $\underline{\boldsymbol{\zeta}}$  and  $\underline{\boldsymbol{\theta}}$ , while for matrix functions  $\underline{\boldsymbol{\zeta}}$  and  $\underline{\boldsymbol{\theta}}$ , the same notation  $(\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\theta}})_D$  denotes the integral over  $D$  of their Frobenius inner product  $\underline{\boldsymbol{\zeta}} : \underline{\boldsymbol{\theta}}$ . Note that in (1.4),  $\underline{\mathbf{L}}^2(\Omega)$  is the space of matrix-valued functions whose entries are in  $L^2(\Omega)$ , and  $\underline{\mathbf{H}}(\text{div}, \Omega)$  is the space of matrix-valued functions whose row vectors belong to the standard  $\mathbf{H}(\text{div}; \Omega)$ -space. Note also that now  $\mathcal{A}$  is extended so that it is symmetric positive definite over all matrices. In Table 1, the space  $\underline{\mathbf{RT}}^k(K)$  denotes the space of all matrices whose row-vectors belong to the standard Raviart-Thomas space

Method	$\underline{\mathbf{V}}(K)$	$\underline{\mathbf{W}}(K)$	$\underline{\mathbf{A}}(K)$	Order
Stenberg [24]	$\underline{\mathcal{P}}^k(K) + \text{curl}(\underline{\mathcal{P}}^{k-1}(K)\lambda_0\lambda_1\lambda_2\lambda_3)$	$\mathcal{P}^{k-1}(K)$	$\underline{\mathbf{A}}^k(K)$	$k \geq 1$
PEERS [4, 12]	$\underline{\mathbf{RT}}^0(K) + \text{curl}(\underline{\mathcal{P}}^0(K)\lambda_0\lambda_1\lambda_2\lambda_3)$	$\mathcal{P}^0(K)$	$C\underline{\mathbf{A}}^1(K)$	$k = 0$
AFW [5]	$\underline{\mathcal{P}}^{k+1}(K)$	$\mathcal{P}^k(K)$	$\underline{\mathbf{A}}^k(K)$	$k \geq 0$
Our 1st element [15]	$\underline{\mathbf{RT}}^k(K) + \text{curl}((\text{curl}\tilde{\underline{\mathbf{A}}}^k(K))\mathbf{b}_K)$	$\mathcal{P}^k(K)$	$\underline{\mathbf{A}}^k(K)$	$k \geq 1$
Section 2 (this paper)	$\underline{\mathcal{P}}^k(K) + \text{curl}((\text{curl}\tilde{\underline{\mathbf{A}}}^k(K))\mathbf{b}_K)$	$\mathcal{P}^{k-1}(K)$	$\underline{\mathbf{A}}^k(K)$	$k \geq 1$
Section 6 (this paper)	$\underline{\mathcal{P}}^k(K)$ , on special grids	$\mathcal{P}^{k-1}(K)$	$\underline{\mathbf{A}}^k(K)$	$k \geq 2$

TABLE 1. Some combinations of stable tetrahedral finite element spaces for mixed methods with weakly imposed stress symmetry. (The prefix  $C$  in  $C\underline{\mathbf{A}}^1(K)$  indicates globally continuous space.)

$\mathcal{P}^k(K) + \mathbf{x}\mathcal{P}^k(K)$  and  $\mathcal{P}^k(K)$  denote vector functions with components in  $\mathcal{P}^k(K)$ . Since the weak symmetry constraint (1.5c) is imposed using polynomials of the same degree for all rows in Table 1 (the space  $\underline{\mathbf{A}}^k(K)$ ) we find that the stress space of Section 2 has the smallest dimension (except for the elements considered in Section 6 which require special meshes for stability). Table 2 gives the dimension counts of our new element and compares it to the two elements most recently discovered, [5] and [15].

The choice of spaces considered in [15] can be viewed as a natural matrix extension of the Raviart-Thomas (RT) [22] element for vector functions, whereas the spaces we consider here could be thought of as the extension of the Brezzi-Douglas-Marini (BDM) [9, 10, 21] element. The choice of the displacement space of one less degree underlines the similarity with the BDM element. However, the AFW element [5] can also be considered as the analogue of the BDM element (and is indeed considered as such in [12]), so we shall refrain from calling our new element ‘the BDM analogue’. There appears to be room for multiple matrix analogues of the vector elements.

The organization of the paper is as follows. In the next section we discuss our new family of elements in two and three dimensions. In Section 3 we define the corresponding projection operator and prove key properties. In Section 4 we give an error analysis of the mixed method. In Section 5 we use connections to the Stokes elements previously made in [2, 12, 17] to put our two-dimensional element in perspective. In the Section 6 we again use the connection with Stokes elements to argue that in special meshes we can drop the bubble term appearing in the definition of  $\underline{\mathbf{V}}(K)$  and yet obtain a stable method which in fact produces stress tensors that are exactly symmetric. In Section 8 we present the hybridized version of our method. In the final section we provide a post-processing technique which gives an improved approximation for the displacement.

Method	Formula	Stress space dimension					
		$k = 0$	1	2	3	4	5
AFW [5]	$p_k + 9(k+2)(k+3)/2$	36	90	180	315	504	756
Our 1st element [15]	$p_k + 3(k+1)(k+2)$	×	54	126	240	405	630
Section 2 (this paper)	$p_k + 3(k+1)(k+2)/2$	×	45	108	210	360	567

TABLE 2. Dimension comparison with two recent works. Here  $p_k$  denotes  $\dim(\underline{\mathcal{P}}^k(K)) = 3(k+1)(k+2)(k+3)/2$ .

## 2. THE NEW FINITE ELEMENT

In this section we define our new finite element for stresses in two and three dimensions. The definition is in the style of Ciarlet [13], whereby a finite element is denoted by a triple describing the geometry, the space, and a set of degrees of freedom. The triple in our case is  $(K, \underline{\mathbf{V}}(K), \Sigma)$ , where the geometric object  $K$  is a simplex (triangle or tetrahedron) and  $\underline{\mathbf{V}}(K)$  is given by

$$\underline{\mathbf{V}}(K) = \underline{\mathcal{P}}^k(K) + B(\underline{\tilde{\mathbf{A}}}^k(K)), \quad (2.6a)$$

$$\underline{\mathbf{W}}(K) = \underline{\mathcal{P}}^{k-1}(K), \quad (2.6b)$$

$$\underline{\mathbf{A}}(K) = \underline{\mathbf{A}}^k(K). \quad (2.6c)$$

where, letting  $d$  denote the ambient space dimension,

$$B(\underline{\boldsymbol{\eta}}) := \begin{cases} \text{curl}(\text{curl}(\underline{\boldsymbol{\eta}}) b_K) & \text{if } d = 2, \\ \text{curl}(\text{curl}(\underline{\boldsymbol{\eta}}) \underline{\mathbf{b}}_K) & \text{if } d = 3, \end{cases}$$

$b_K$  denotes the (scalar) bubble of the simplex  $K$  equaling the product of all its barycentric coordinates (so when  $d = 2$ , we have  $b_K = \lambda_0 \lambda_1 \lambda_2$ ), and  $\underline{\mathbf{b}}_K$  is the symmetric matrix bubble defined by (1.2). As mentioned above, the operator curl when applied to a matrix acts row-wise in three dimensions. In two dimensions, curl applied to a matrix and a vector is defined in the following way

$$\text{curl}(\underline{\boldsymbol{\eta}}) := \begin{pmatrix} \partial_1 \eta_{12} - \partial_2 \eta_{11} \\ \partial_1 \eta_{22} - \partial_2 \eta_{21} \end{pmatrix} \quad \text{and} \quad \text{curl}(\underline{\mathbf{w}}) := \begin{pmatrix} \partial_2 w_1 & -\partial_1 w_1 \\ \partial_2 w_2 & -\partial_1 w_2 \end{pmatrix},$$

respectively.

The set of degrees of freedom  $\Sigma$  are the following linear functionals

$$\ell_{\boldsymbol{\mu}}(\underline{\boldsymbol{\sigma}}) = \langle \underline{\boldsymbol{\sigma}} \mathbf{n}, \boldsymbol{\mu} \rangle_F, \quad \text{for all } \boldsymbol{\mu} \in \underline{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \quad (2.7a)$$

$$\ell_{\underline{\mathbf{v}}}(\underline{\boldsymbol{\sigma}}) = (\underline{\boldsymbol{\sigma}}, \underline{\mathbf{v}})_K, \quad \text{for all } \underline{\mathbf{v}} \in \underline{\mathbf{N}}^{k-1}(K), \quad (2.7b)$$

$$\ell_{\underline{\boldsymbol{\eta}}}(\underline{\boldsymbol{\sigma}}) = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\tilde{\mathbf{A}}}^k(K). \quad (2.7c)$$

Here  $\underline{\mathbf{N}}^s(K)$  is the space of matrices whose row-vectors are in the Nédélec space of the first kind [20] of index  $s$ , i.e., it is defined by

$$\underline{\mathbf{N}}^s(K) := \underline{\mathcal{P}}^{s-1}(K) + \underline{\mathbf{S}}^s(K),$$

where

$$\underline{\mathbf{S}}^s(K) := \{\underline{\mathbf{v}} \in \underline{\tilde{\mathcal{P}}}^s(K) : \underline{\mathbf{v}} \mathbf{x} = 0\},$$

for  $s \geq 1$ . Here  $\underline{\tilde{\mathcal{P}}}^s(K) = [\tilde{\mathcal{P}}^s(K)]^{d \times d}$  and  $\tilde{\mathcal{P}}^s(K)$  are the homogeneous polynomials of degree  $s$ . For  $s = 0$ ,  $\underline{\mathbf{N}}^0(K)$  is the empty set. The main result of this section is the following theorem.

**Theorem 2.1** (Unisolvency). *Assume  $k \geq 1$ . Let  $K$  be a triangle ( $d=2$ ) or a tetrahedron ( $d=3$ ). Any  $\underline{\boldsymbol{\sigma}}$  in  $\underline{\mathbf{V}}^k(K)$  is uniquely determined by the degrees of freedom given by (2.7).*

Before proving this theorem we need to recall a few results we proved in [15], which are summarized in the next lemma. Here onwards we will assume  $k \geq 1$  tacitly.

**Lemma 2.2.** *The following statements hold:*

(1) The form given by

$$((\underline{\mathbf{u}}, \underline{\mathbf{v}}))_b := \begin{cases} (\underline{\mathbf{u}} \mathbf{b}_K, \underline{\mathbf{v}})_K & \text{if } d = 2, \\ (\underline{\mathbf{u}} \underline{\mathbf{b}}_K, \underline{\mathbf{v}})_K & \text{if } d = 3, \end{cases}$$

is an inner product on  $\underline{\mathcal{P}}^k(K)$ .

- (2) For any matrix  $\underline{\mathbf{u}}$ , all the row vectors of the products  $\underline{\mathbf{u}} \underline{\mathbf{b}}_K$  and  $\underline{\mathbf{u}} \mathbf{b}_K$  have vanishing tangential trace on  $\partial K$ .
- (3) If  $B(\underline{\boldsymbol{\eta}}) = \underline{\boldsymbol{\psi}}$  with  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  and  $\underline{\boldsymbol{\eta}} \in \tilde{\underline{\mathcal{A}}}^k(K)$  then  $\underline{\boldsymbol{\eta}} = 0$ .
- (4) Any  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  satisfies  $\operatorname{div} \underline{\boldsymbol{\psi}} = 0$  and  $\underline{\boldsymbol{\psi}} \mathbf{n}|_{\partial K} = 0$  if and only if there exists a  $\underline{\mathbf{v}} \in \underline{\mathcal{P}}^{k-1}(K)$  such that  $\underline{\boldsymbol{\psi}} = B(\underline{\mathbf{v}})$ .

*Proof.* Items (1)–(3) are proved in [15]. Item (4), with  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  replaced by  $\underline{\boldsymbol{\psi}} \in \underline{\mathbf{RT}}^k(K)$ , is also proved in [15]. That it follows with  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  is a consequence of the well known fact that the divergence free subspaces of  $\underline{\mathbf{RT}}^k(K)$  and  $\underline{\mathcal{P}}^k(K)$  coincide.  $\square$

We also need the following simple corollary of Lemma 2.2(4).

**Lemma 2.3.** Any  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  satisfies  $\operatorname{div} \underline{\boldsymbol{\psi}} = 0$  and  $\underline{\boldsymbol{\psi}} \mathbf{n}|_{\partial K} = 0$  if and only if there exists a Nédélec polynomial  $\underline{\mathbf{v}} \in \underline{\mathbf{N}}^{k-1}(K)$  such that  $\underline{\boldsymbol{\psi}} = B(\underline{\mathbf{v}})$ .

*Proof.* From Lemma 2.2 it is obvious that if  $\underline{\boldsymbol{\psi}} = B(\underline{\mathbf{v}})$  for some  $\underline{\mathbf{v}}$  in  $\underline{\mathbf{N}}^{k-1}(K)$ , then  $\operatorname{div} \underline{\boldsymbol{\psi}} = 0$  and  $\underline{\boldsymbol{\psi}} \mathbf{n}|_{\partial K} = 0$ .

For the converse, consider a  $\underline{\boldsymbol{\psi}}$  satisfying  $\operatorname{div} \underline{\boldsymbol{\psi}} = 0$  and  $\underline{\boldsymbol{\psi}} \mathbf{n}|_{\partial K} = 0$ . By Lemma 2.2(4), we know that there is a  $\underline{\mathbf{v}}$  in  $\underline{\mathcal{P}}^{k-1}(K)$  such that  $\underline{\boldsymbol{\psi}} = B(\underline{\mathbf{v}})$ . Decompose  $\underline{\mathbf{v}}$  as

$$\underline{\mathbf{v}} = \underline{\boldsymbol{\zeta}} + \underline{\boldsymbol{\tau}},$$

where  $\underline{\boldsymbol{\zeta}} \in \underline{\mathcal{P}}^{k-2}(K)$  and  $\underline{\boldsymbol{\tau}} \in \tilde{\underline{\mathcal{P}}}^{k-1}(K)$ . Let  $\underline{\mathbf{z}} = -\underline{\boldsymbol{\tau}} \mathbf{x}/k$ , and define

$$\underline{\mathbf{w}} = \underline{\mathbf{v}} + \operatorname{grad} \underline{\mathbf{z}}.$$

Then, we see that  $\operatorname{curl}(\underline{\boldsymbol{\psi}}) = \operatorname{curl}(\underline{\mathbf{w}})$  and so

$$\underline{\boldsymbol{\psi}} = B(\underline{\mathbf{w}}).$$

To complete the proof, we will show that  $\underline{\mathbf{w}} \in \underline{\mathbf{N}}^{k-1}(K)$ . Note that  $\underline{\mathbf{w}} = \underline{\boldsymbol{\zeta}} + \underline{\boldsymbol{\tau}} + \operatorname{grad} \underline{\mathbf{z}}$ , where  $\underline{\boldsymbol{\zeta}} \in \underline{\mathcal{P}}^{k-2}(K)$  and  $\underline{\boldsymbol{\tau}} + \operatorname{grad} \underline{\mathbf{z}} \in \tilde{\underline{\mathcal{P}}}^{k-1}(K)$ . A simple calculation (Euler's identity) shows that due to the homogeneity of  $\underline{\boldsymbol{\tau}}$ , one has  $\operatorname{grad}(\underline{\boldsymbol{\tau}} \mathbf{x}) \mathbf{x} = k \underline{\boldsymbol{\tau}} \mathbf{x}$ . Therefore,

$$\begin{aligned} (\underline{\boldsymbol{\tau}} + \operatorname{grad} \underline{\mathbf{z}}) \mathbf{x} &= \underline{\boldsymbol{\tau}} \mathbf{x} + (\operatorname{grad} \underline{\mathbf{z}}) \mathbf{x} = \underline{\boldsymbol{\tau}} \mathbf{x} - \left( \frac{1}{k} \operatorname{grad}(\underline{\boldsymbol{\tau}} \mathbf{x}) \right) \mathbf{x} \\ &= \underline{\boldsymbol{\tau}} \mathbf{x} - \underline{\boldsymbol{\tau}} \mathbf{x} = 0. \end{aligned}$$

This establishes that  $\underline{\boldsymbol{\tau}} + \operatorname{grad} \underline{\mathbf{z}}$  is in  $\underline{\mathbf{S}}^{k-1}(K)$  and hence  $\underline{\mathbf{w}} \in \underline{\mathbf{N}}^{k-1}(K)$ .  $\square$

Now we can prove the main result of this section.

*Proof of Theorem 2.1.* We first show that  $\dim(\underline{\mathbf{V}}(K))$  is equal to the number of independent degrees of freedom given in (2.7). To this end, note that

$$\dim(\underline{\mathbf{V}}(K)) = \dim(\underline{\mathcal{P}}^k(K)) + \dim(\tilde{\underline{\mathcal{A}}}(K)),$$

where we used Lemma 2.2(3). Hence, we must show that

$$\dim(\underline{\mathcal{P}}^k(k)) = \dim(\underline{\mathbf{N}}^{k-1}(K)) + (d+1) \dim \underline{\mathcal{P}}^k(F), \quad (2.8)$$

where  $F$  is a face of  $K$ . The dimension of  $\underline{\mathbf{N}}^{k-1}(K)$  can be calculated (see e.g. [21] for  $d = 3$  or [18] for general  $d$ ) to be

$$\dim(\underline{\mathbf{N}}^{k-1}(K)) = \frac{(k-1+d)!(k-1)}{(d-1)!k!}d.$$

Hence the right hand side of (2.8) equals

$$\frac{(k-1+d)!(k-1)}{(d-1)!k!}d + (d+1)\frac{(k-1+d)!}{(d-1)!k!}d = \frac{(k+d)!}{k!d!}d^2 = \dim(\underline{\mathcal{P}}^k(K))$$

which proves (2.8).

We next show that if all the degrees of freedom in (2.7) applied to a  $\underline{\boldsymbol{\sigma}}$  in  $\underline{\mathbf{V}}(K)$  vanish, then  $\underline{\boldsymbol{\sigma}}$  must vanish. Since  $\underline{\boldsymbol{\sigma}} \in \underline{\mathbf{V}}(K)$ , by definition, we can write  $\underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\psi}} + B(\underline{\boldsymbol{\zeta}})$  for some  $\underline{\boldsymbol{\psi}} \in \underline{\mathcal{P}}^k(K)$  and  $\underline{\boldsymbol{\zeta}} \in \underline{\tilde{\mathbf{A}}}(K)$ . Standard arguments using  $\ell_{\boldsymbol{\mu}}(\underline{\boldsymbol{\sigma}}) = \ell_{\underline{\boldsymbol{v}}}(\underline{\boldsymbol{\sigma}}) = 0$  show (see, e.g. [15]) that

$$\operatorname{div} \underline{\boldsymbol{\psi}} = 0 \text{ on } K \quad \text{and} \quad \underline{\boldsymbol{\psi}} \mathbf{n} = 0 \text{ on } \partial K. \quad (2.9)$$

Hence, by Lemma 2.3, we have  $\underline{\boldsymbol{\sigma}} = B(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}})$  for some  $\underline{\boldsymbol{w}} \in \underline{\mathbf{N}}^{k-1}(K)$  and  $\underline{\boldsymbol{\zeta}} \in \underline{\tilde{\mathbf{A}}}(K)$ . Furthermore, since  $\ell_{\underline{\boldsymbol{v}}}(\underline{\boldsymbol{\sigma}}) = \ell_{\underline{\boldsymbol{\eta}}}(\underline{\boldsymbol{\sigma}}) = 0$ , we have

$$0 = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}})_K = (B(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}}), \underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}})_K = ((\operatorname{curl}(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}}), \operatorname{curl}(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}})))_b,$$

where we have used integration by parts and Lemma 2.2(2). Using Lemma 2.2(1) also, we conclude that  $\operatorname{curl}(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}}) = 0$  and therefore  $\underline{\boldsymbol{\sigma}} = B(\underline{\boldsymbol{w}} + \underline{\boldsymbol{\zeta}}) = 0$ . This completes the proof.  $\square$

### 3. THE PROJECTION

The natural interpolant of our finite element defines a Fortin projector with a commutativity property involving the row-wise divergence. However, unlike the analogous projector of our first element [15], the new projector does not satisfy the weak symmetry condition in addition to the commutativity property. Specifically, if  $\underline{\boldsymbol{\Pi}}^{(1)}$  denotes the projection of [15], then

$$(\underline{\boldsymbol{\Pi}}^{(1)} \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^k(K),$$

so that if  $\underline{\boldsymbol{\sigma}}$  is symmetric, then  $\underline{\boldsymbol{\Pi}}^{(1)} \underline{\boldsymbol{\sigma}}$  is weakly symmetric. But for the current element, we are only able to show that the above holds under an additional condition (see the next theorem).

Our projector is denoted by  $\underline{\boldsymbol{\Pi}}$  and is defined by  $\ell(\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}) = \ell(\underline{\boldsymbol{\sigma}})$  for all  $\ell$  in  $\Sigma$ , i.e.,  $\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}$  is the unique function in  $\underline{\mathbf{V}}(K)$  satisfying

$$\langle (\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}) \mathbf{n}, \underline{\boldsymbol{\mu}} \rangle_F = \langle \underline{\boldsymbol{\sigma}} \mathbf{n}, \underline{\boldsymbol{\mu}} \rangle_F, \quad \text{for all } \underline{\boldsymbol{\mu}} \in \underline{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \quad (3.10a)$$

$$(\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_K = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_K, \quad \text{for all } \underline{\boldsymbol{v}} \in \underline{\mathbf{N}}^{k-1}(K), \quad (3.10b)$$

$$(\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\tilde{\mathbf{A}}}^k(K). \quad (3.10c)$$

Clearly, by the choice of the degrees of freedom,  $\underline{\boldsymbol{\Pi}} \underline{\boldsymbol{\sigma}}$  is in  $\underline{\mathbf{H}}(\operatorname{div}, \Omega)$ . The domain of  $\underline{\boldsymbol{\Pi}}$  consist of functions  $\underline{\boldsymbol{\sigma}}$  in  $\underline{\mathbf{H}}(\operatorname{div}, \Omega)$  which are regular enough for all the degrees of freedom  $\ell_{\boldsymbol{\mu}}(\underline{\boldsymbol{\sigma}})$  to exist, e.g.,  $\underline{\boldsymbol{\sigma}} \in \underline{\mathbf{H}}(\operatorname{div}, \Omega) \cap \underline{\mathbf{L}}^p(\Omega)$  for  $p > 2$ . The following theorem lists all the properties of the projection we shall need.

**Theorem 3.1.** *Let  $\mathbf{P}$  be the  $L^2(\Omega)$ -orthogonal projection onto  $\mathbf{W}^h$  and let  $\underline{\boldsymbol{\sigma}} \in \operatorname{dom}(\underline{\boldsymbol{\Pi}})$ .*

(1) The following commutativity property holds:

$$\operatorname{div}(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}}) = \mathbf{P}(\operatorname{div}\underline{\boldsymbol{\sigma}}). \quad (3.11)$$

(2) If  $\underline{\boldsymbol{\sigma}}$  is such that  $\operatorname{div}\underline{\boldsymbol{\sigma}}$  is in  $\mathbf{W}_h$ , then

$$(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^k(K).$$

(3) If  $\underline{\boldsymbol{\sigma}}$  is in  $\underline{\mathbf{H}}^r(K)$  for some  $r$  satisfying  $1 \leq r \leq k+1$ , then

$$\|\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(K)} \leq C h_K^r |\underline{\boldsymbol{\sigma}}|_{\underline{\mathbf{H}}^r(K)}. \quad (3.12)$$

where  $h_K = \operatorname{diam}(K)$ .

In the remainder of this section, we prove this theorem. In (3.12) and elsewhere, we use  $C$  to denote a generic constant dependent only on the shape regularity of the elements (and not on  $h_K$ ). Let us begin by proving the first assertion of the theorem.

*Proof of Theorem 3.1(1).* To prove the identity (3.11), first observe that for any  $\boldsymbol{\omega}$  in  $\mathcal{P}^{k-1}(K)$  we have that  $\operatorname{grad}\boldsymbol{\omega} \in \underline{\mathbf{N}}^{k-1}(K)$ . Therefore,

$$\begin{aligned} (\operatorname{div}(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}}), \boldsymbol{\omega})_K &= -(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}}, \operatorname{grad}\boldsymbol{\omega})_K + \langle \underline{\mathbf{P}}\underline{\boldsymbol{\sigma}}\mathbf{n}, \boldsymbol{\omega} \rangle_{\partial K} \\ &= -(\underline{\boldsymbol{\sigma}}, \operatorname{grad}\boldsymbol{\omega})_K + \langle \underline{\boldsymbol{\sigma}}\mathbf{n}, \boldsymbol{\omega} \rangle_{\partial K} && \text{by (3.10b) and (3.10a),} \\ &= (\operatorname{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{\omega})_K, && \text{after integration by parts,} \\ &= (\mathbf{P}\operatorname{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{\omega})_K. \end{aligned}$$

This proves item (1) of Theorem 3.1.  $\square$

*Proof of Theorem 3.1(2).* Let  $\underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^k(K)$ . We know that  $\underline{\boldsymbol{\eta}}$  can be written in the following form

$$\underline{\boldsymbol{\eta}} = \underline{\mathbf{v}} + \underline{\tilde{\boldsymbol{\eta}}}.$$

where  $\underline{\mathbf{v}} \in \underline{\mathcal{P}}^{k-1}(K)$  and  $\underline{\tilde{\boldsymbol{\eta}}} \in \underline{\tilde{\mathbf{A}}}^k(K)$ . Moreover, from [20, 21], we know that

$$\underline{\mathbf{v}} = \underline{\boldsymbol{\psi}} + \operatorname{grad}\boldsymbol{\omega}$$

for some  $\underline{\boldsymbol{\psi}} \in \underline{\mathbf{N}}^{k-1}(K)$  and  $\boldsymbol{\omega} \in \tilde{\mathcal{P}}^k(K) \equiv [\tilde{\mathcal{P}}^k(K)]^d$ . Hence, using (3.10b), (3.10c) we have

$$(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K = (\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}, \operatorname{grad}\boldsymbol{\omega})_K.$$

Integrating by parts and using (3.10a), we get

$$(\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K = -(\operatorname{div}\underline{\mathbf{P}}\underline{\boldsymbol{\sigma}} - \operatorname{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{\omega})_K = -(\mathbf{P}\operatorname{div}\underline{\boldsymbol{\sigma}} - \operatorname{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{\omega})_K.$$

where we used (3.11). This completes the proof.  $\square$

It only remains to prove item (3) of Theorem 3.1. This can be done by a Bramble-Hilbert argument once we make the mappings clear. Let  $\hat{K}$  be a fixed ‘reference’ simplex. It can be mapped homeomorphically to any other simplex  $K$  by a mapping of the form  $\mathcal{F}(\hat{\mathbf{x}}) = \underline{\mathbf{M}}\hat{\mathbf{x}} + \mathbf{b}$  where  $\underline{\mathbf{M}}$  is a constant matrix and  $\mathbf{b}$  a constant vector. We use the same mappings introduced in [15]. Matrix-valued functions  $\hat{\boldsymbol{\sigma}}$  and (column) vector-valued functions  $\hat{\boldsymbol{w}}$  on  $\hat{K}$  are mapped to functions on  $K$  by

$$\boldsymbol{\sigma}(\mathbf{x}) = |\det \underline{\mathbf{M}}|^{-1} \underline{\mathbf{M}} \hat{\boldsymbol{\sigma}}(\hat{\mathbf{x}}) \underline{\mathbf{M}}^t, \quad (3.13a)$$

$$\boldsymbol{w}(\mathbf{x}) = \underline{\mathbf{M}}^{-t} \hat{\boldsymbol{w}}(\hat{\mathbf{x}}), \quad (3.13b)$$

respectively, and let

$$\check{\boldsymbol{\tau}}(\hat{\boldsymbol{x}}) = \underline{\boldsymbol{M}}^t \boldsymbol{\tau}(\boldsymbol{x}) \underline{\boldsymbol{M}}. \quad (3.13c)$$

**Lemma 3.2.** *The mapping  $\boldsymbol{\tau} \mapsto \check{\boldsymbol{\tau}}$  is a homeomorphism from  $\underline{\boldsymbol{N}}^k(K)$  onto  $\underline{\boldsymbol{N}}^k(\hat{K})$ .*

*Proof.* It suffices to prove that if  $\boldsymbol{\tau}$  is in  $\underline{\boldsymbol{N}}^k(K)$ , then  $\check{\boldsymbol{\tau}}$  is in  $\underline{\boldsymbol{N}}^k(\hat{K})$ . To this end, we make two observations. First, note that

$$\begin{aligned} \check{\boldsymbol{\tau}} \hat{\boldsymbol{x}} &= (\underline{\boldsymbol{M}}^t \boldsymbol{\tau}(\boldsymbol{x}) \underline{\boldsymbol{M}}) \hat{\boldsymbol{x}} = (\underline{\boldsymbol{M}}^t \boldsymbol{\tau}(\boldsymbol{x}) \underline{\boldsymbol{M}}) \underline{\boldsymbol{M}}^{-1}(\boldsymbol{x} - \boldsymbol{b}) \\ &= -\underline{\boldsymbol{M}}^t \boldsymbol{\tau}(\boldsymbol{x}) \boldsymbol{b} + \underline{\boldsymbol{M}}^t \boldsymbol{\tau}(\boldsymbol{x}) \boldsymbol{x}. \end{aligned} \quad (3.14)$$

The second observation is that a function  $\boldsymbol{r}$  is in  $\underline{\boldsymbol{N}}^k(K)$  if and only if  $\boldsymbol{r}$  is in  $\underline{\boldsymbol{\mathcal{P}}}^k(K)$  and the product  $\boldsymbol{r} \boldsymbol{x}$  is in  $\boldsymbol{\mathcal{P}}^k(K)$ . This implies that the last term in (3.14) is in  $\boldsymbol{\mathcal{P}}^k(K)$  (as  $\boldsymbol{\tau}$  is in  $\underline{\boldsymbol{N}}^k(K)$ ). Since the penultimate term in (3.14) is obviously also in  $\underline{\boldsymbol{N}}^k(K)$ , we find that  $\check{\boldsymbol{\tau}} \hat{\boldsymbol{x}}$  is in  $\boldsymbol{\mathcal{P}}^k(\hat{K})$ , so  $\check{\boldsymbol{\tau}}$  is in  $\underline{\boldsymbol{N}}^k(\hat{K})$ .  $\square$

*Proof of Theorem 3.1(3).* Let  $\hat{\boldsymbol{\Pi}}$  denote the interpolant of the reference element and let  $\underline{\boldsymbol{\sigma}}$  and  $\hat{\boldsymbol{\sigma}}$  be as in (3.13a). Let us first prove that

$$\widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}} = \hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\sigma}}, \quad (3.15)$$

where  $\widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}}$  is the function obtained by mapping  $\boldsymbol{\Pi}\boldsymbol{\sigma}$  from  $K$  to  $\hat{K}$  using the map in (3.13a). Simple calculations show that for all  $\boldsymbol{r}$  in  $\underline{\boldsymbol{N}}^k(K)$ ,

$$\begin{aligned} (\widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}}, \check{\boldsymbol{r}})_{\hat{K}} &= (\boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{r})_K = (\boldsymbol{\sigma}, \boldsymbol{r})_K = (\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{r}})_{\hat{K}} \\ &= (\hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{r}})_{\hat{K}}, \end{aligned}$$

where we have used (3.10b) on  $K$  as well as  $\hat{K}$ , and Lemma 3.2. Combining with similar identities derived using the other degrees of freedom, and the properties of the mappings in (3.13), we have

$$\begin{aligned} (\widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}}, \check{\boldsymbol{r}})_{\hat{K}} &= (\hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{r}})_{\hat{K}}, & \text{for all } \check{\boldsymbol{r}} \in \underline{\boldsymbol{N}}^k(\hat{K}), \\ (\widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}}, \check{\boldsymbol{\eta}})_{\hat{K}} &= (\hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{\eta}})_{\hat{K}}, & \text{for all } \check{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}^k(\hat{K}), \\ \langle \widehat{\boldsymbol{\Pi}\boldsymbol{\sigma}\boldsymbol{n}}, \hat{\boldsymbol{\mu}} \rangle_{\partial\hat{K}} &= \langle \hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\sigma}\boldsymbol{n}}, \hat{\boldsymbol{\mu}} \rangle_{\partial\hat{K}} & \text{for all } \hat{\boldsymbol{\mu}} \in \boldsymbol{\mathcal{P}}^k(\hat{F}), \text{ for all faces } \hat{F} \text{ of } \hat{K}. \end{aligned}$$

Hence, by the unisolvency result of Theorem 2.1, we conclude (3.15).

The rest of the proof can now proceed by a standard Bramble-Hilbert argument (we omit the details).  $\square$

*Remark 3.3.* To connect our projection with the standard ‘‘BDM projection’’ as defined in [21], consider the obvious generalization of the projection of [21] defined (for matrices in both two and three dimensions) by

$$\langle (\underline{\boldsymbol{\Pi}}_{\text{BDM}}\boldsymbol{\sigma})\boldsymbol{n}, \boldsymbol{\mu} \rangle_F = \langle \boldsymbol{\sigma}\boldsymbol{n}, \boldsymbol{\mu} \rangle_F, \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \quad (3.16a)$$

$$(\underline{\boldsymbol{\Pi}}_{\text{BDM}}\boldsymbol{\sigma}, \boldsymbol{v})_K = (\boldsymbol{\sigma}, \boldsymbol{v})_K, \quad \text{for all } \boldsymbol{v} \in \underline{\boldsymbol{N}}^{k-1}(K). \quad (3.16b)$$

Comparing with (3.10), we find that  $\boldsymbol{\Pi}\boldsymbol{\sigma} - \underline{\boldsymbol{\Pi}}_{\text{BDM}}\boldsymbol{\sigma}$  is a function for which all our degrees of freedom vanish except those of the type (2.7c).



## 4. ERROR ANALYSIS

In this section we return to the mixed method (1.5) with the specific choices of our local finite element spaces. We will prove that the method is stable and optimally convergent. The error analysis of this section proceeds by obtaining first results as an application of the Babuška-Brezzi theory [11] and then refining the estimates to separate the errors in each variable. In either case, the projection  $\underline{\mathbf{I}}$  plays an important role.

To apply the Babuška-Brezzi theory, we must verify (i) the ‘‘inf-sup condition’’

$$\sup_{\underline{\boldsymbol{\tau}} \in \underline{\mathbf{V}}^h} \frac{(\mathbf{w}^h, \operatorname{div} \underline{\boldsymbol{\tau}}) + (\boldsymbol{\zeta}^h, \underline{\boldsymbol{\tau}})}{\|\underline{\boldsymbol{\tau}}\|_{H(\operatorname{div}, \Omega)}} \geq C(\|\mathbf{w}^h\|_{L^2(\Omega)} + \|\boldsymbol{\zeta}^h\|_{L^2(\Omega)}) \quad (4.17)$$

for all  $\boldsymbol{\zeta}^h$  in  $\underline{\mathbf{A}}^h$  and  $\mathbf{w}^h$  in  $\mathbf{W}^h$ , and (ii) the ‘‘coercivity in the kernel’’ condition

$$(\mathcal{A}\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\tau}})_\Omega \geq C\|\underline{\boldsymbol{\tau}}\|_{H(\operatorname{div}, \Omega)} \quad (4.18)$$

for all  $\underline{\boldsymbol{\tau}}$  in the ‘kernel’ given by  $\{\underline{\boldsymbol{\tau}} \in \underline{\mathbf{V}}^h : (\mathbf{w}^h, \operatorname{div} \underline{\boldsymbol{\tau}})_\Omega + (\boldsymbol{\eta}^h, \underline{\boldsymbol{\tau}})_\Omega = 0 \text{ for all } \mathbf{w}^h \in \mathbf{W}^h \text{ and all } \boldsymbol{\eta}^h \in \underline{\mathbf{A}}^h\}$ . Due to the surjectivity of  $\operatorname{div} : \underline{\mathbf{V}}^h \mapsto \mathbf{W}^h$ , condition (4.18) is obvious in our case. To establish (4.17), we use the following lemma.

**Lemma 4.1.** *Given any  $\mathbf{w} \in \mathbf{W}^h$  and  $\boldsymbol{\zeta} \in \underline{\mathbf{A}}^h$ , there exists a  $\underline{\boldsymbol{\tau}} \in \underline{\mathbf{V}}^h$  satisfying*

$$\operatorname{div} \underline{\boldsymbol{\tau}} = \mathbf{w}, \quad (4.19a)$$

$$(\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\eta}}) = (\boldsymbol{\zeta}, \underline{\boldsymbol{\eta}}), \quad \forall \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^h, \text{ and} \quad (4.19b)$$

$$\|\underline{\boldsymbol{\tau}}\|_{L^2(\Omega)} \leq C(\|\mathbf{w}\|_{L^2(\Omega)} + \|\boldsymbol{\zeta}\|_{L^2(\Omega)}), \quad (4.19c)$$

where  $C$  only depends on the shape regularity of the mesh.

*Proof.* By [5, Theorem 7.1], there exists a  $\underline{\boldsymbol{\tau}}_{k+1}$  in the space  $\{\mathbf{v} \in \underline{\mathbf{H}}(\operatorname{div}, \Omega) : \mathbf{v}|_K \in \underline{\mathcal{P}}^{k+1}(K), \text{ for all } K \in \Omega_h\}$  such that

$$\operatorname{div} \underline{\boldsymbol{\tau}}_{k+1} = \mathbf{w},$$

$$(\underline{\boldsymbol{\tau}}_{k+1}, \underline{\boldsymbol{\eta}}) = (\boldsymbol{\zeta}, \underline{\boldsymbol{\eta}}), \quad \forall \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^h, \text{ and}$$

$$\|\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(\Omega)} \leq C(\|\mathbf{w}\|_{L^2(\Omega)} + \|\boldsymbol{\zeta}\|_{L^2(\Omega)}), \quad (4.20)$$

where  $C$  depends on the shape regularity of the mesh. Being a piecewise polynomial,  $\underline{\boldsymbol{\tau}}_{k+1}$  obviously has enough regularity for applying  $\underline{\mathbf{I}}$ , so we may set  $\underline{\boldsymbol{\tau}} = \underline{\mathbf{I}}\underline{\boldsymbol{\tau}}_{k+1}$ . Then

$$\operatorname{div} \underline{\boldsymbol{\tau}} = \mathbf{P} \operatorname{div} \underline{\boldsymbol{\tau}}_{k+1} = \mathbf{w}, \quad \text{by Theorem 3.1(1)}$$

$$(\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\eta}}) = (\underline{\boldsymbol{\tau}}_{k+1}, \underline{\boldsymbol{\eta}}) = (\boldsymbol{\zeta}, \underline{\boldsymbol{\eta}}), \quad \forall \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^h, \quad \text{by Theorem 3.1(2),}$$

where we used that  $\mathbf{P} \operatorname{div} \underline{\boldsymbol{\tau}}_{k+1} = \mathbf{w} = \operatorname{div} \underline{\boldsymbol{\tau}}_{k+1}$ . This proves the first two equations of the lemma. For the final inequality, we use the triangle inequality and (3.12) to get

$$\begin{aligned} \|\underline{\boldsymbol{\tau}}\|_{L^2(K)} &= \|\underline{\mathbf{I}}\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(K)} \\ &\leq \|\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(K)} + \|\underline{\mathbf{I}}\underline{\boldsymbol{\tau}}_{k+1} - \underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(K)} \\ &\leq \|\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(K)} + Ch_K \|\underline{\boldsymbol{\tau}}_{k+1}\|_{H^1(K)} \\ &\leq C\|\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(K)}, \end{aligned}$$

where we used an inverse estimate since  $\underline{\boldsymbol{\tau}}_{k+1}|_K \in \underline{\mathcal{P}}^{k+1}(K)$ . Using (4.20), we complete the proof.  $\square$

Our first theorem shows that the finite element error can be bounded in terms of the best approximation error:

**Theorem 4.2** (Quasioptimality). *Let  $(\underline{\sigma}, \mathbf{u}, \underline{\rho})$  be the exact solution of (1.3). Suppose  $(\underline{\sigma}^h, \mathbf{u}^h, \underline{\rho}^h)$  satisfies (1.5). Then there is a  $C > 0$  such that*

$$\|\underline{\sigma} - \underline{\sigma}_h\|_{H(\text{div})} + \|\mathbf{u} - \mathbf{u}_h\|_{L^2} + \|\underline{\rho} - \underline{\rho}^h\|_{L^2} \leq C(\|\underline{\sigma} - \underline{\mathbf{v}}\|_{H(\text{div})} + \|\mathbf{u} - \underline{\omega}\|_{L^2} + \|\underline{\rho} - \underline{\eta}\|_{L^2})$$

for all  $(\underline{\mathbf{v}}, \underline{\omega}, \underline{\eta}) \in \underline{\mathbf{V}}^h \times \mathbf{W}^h \times \underline{\mathbf{A}}^h$ . In particular, this implies that the method (1.5) is uniquely solvable.

*Proof.* The inf-sup condition (4.17) is an immediate consequence of Lemma 4.1. Hence, the inequality of the theorem follows from the Babuška-Brezzi theory [11].

Since the exact solution  $(\underline{\sigma}, \mathbf{u}, \underline{\rho})$  is trivial if all the right hand sides vanish, the unique solvability of the method follows from the inequality of the theorem.  $\square$

Since the space where the displacement is approximated is one order less, let us now separate the error in  $\mathbf{u}$  from the remaining variables to get better estimates.

**Theorem 4.3.** *Let  $\underline{\mathbf{P}}$  denote the  $L^2$ -orthogonal projection into  $\underline{\mathbf{A}}^h$ . Then*

$$\|\underline{\sigma} - \underline{\sigma}^h\|_{L^2(\Omega)} + \|\underline{\rho} - \underline{\rho}^h\|_{L^2(\Omega)} \leq C(\|\underline{\Pi}\underline{\sigma} - \underline{\sigma}\|_{L^2(\Omega)} + \|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}\|_{L^2(\Omega)}).$$

To prove this theorem, we use the following lemma which estimates the  $\underline{\rho}$ -error in terms of the error in  $\underline{\sigma}$ .

**Lemma 4.4.** *We have,*

$$\|\underline{\rho} - \underline{\rho}^h\|_{L^2(\Omega)} \leq C(\|\underline{\sigma} - \underline{\sigma}^h\|_{L^2(\Omega)} + \|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}\|_{L^2(\Omega)}).$$

*Proof.* We start by writing the error equations

$$(\mathcal{A}(\underline{\sigma} - \underline{\sigma}^h), \underline{\mathbf{v}})_\Omega + (\mathbf{u} - \mathbf{u}^h, \text{div } \underline{\mathbf{v}})_\Omega + (\underline{\rho} - \underline{\rho}^h, \underline{\mathbf{v}})_\Omega = 0 \quad (4.21a)$$

$$(\text{div } (\underline{\sigma} - \underline{\sigma}^h), \underline{\omega})_\Omega = 0 \quad (4.21b)$$

$$(\underline{\sigma} - \underline{\sigma}^h, \underline{\eta})_\Omega = 0, \quad (4.21c)$$

for all  $(\underline{\mathbf{v}}, \underline{\omega}, \underline{\eta}) \in \underline{\mathbf{V}}^h \times \mathbf{W}^h \times \underline{\mathbf{A}}^h$ . In (4.21a), we set in place of  $\underline{\mathbf{v}}$ , the  $\underline{\boldsymbol{\tau}}$  obtained from Lemma 4.1 with  $\underline{\boldsymbol{\zeta}} = \underline{\mathbf{P}}\underline{\rho} - \underline{\rho}^h$  and  $\mathbf{w} = \text{div } \underline{\boldsymbol{\tau}} = 0$ , so that

$$\begin{aligned} \|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}^h\|_{L^2(\Omega)}^2 &= (\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}^h, \underline{\boldsymbol{\tau}}) \\ &= -(\mathcal{A}(\underline{\sigma} - \underline{\sigma}^h), \underline{\boldsymbol{\tau}}) + (\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}, \underline{\boldsymbol{\tau}}). \end{aligned}$$

Thus,

$$\|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}^h\|_{L^2(\Omega)}^2 \leq (\|\underline{\sigma} - \underline{\sigma}^h\|_{L^2(\Omega)} + \|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}\|_{L^2(\Omega)}) \|\underline{\boldsymbol{\tau}}\|_{L^2(\Omega)}.$$

The proof is complete once we note that  $\|\underline{\boldsymbol{\tau}}\|_{L^2(\Omega)} \leq \|\underline{\mathbf{P}}\underline{\rho} - \underline{\rho}^h\|_{L^2(\Omega)}$  by Lemma 4.1.  $\square$

*Proof of Theorem 4.3.* We again use the error equations (4.21). First note that using (3.11) and (4.21b) we get that

$$\text{div } (\underline{\Pi}\underline{\sigma} - \underline{\sigma}^h) = 0. \quad (4.22)$$

Therefore, using (4.21a) with  $\underline{\mathbf{v}} = \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$  we get

$$\begin{aligned} & (\mathcal{A}(\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h), \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega \\ &= (\mathcal{A}(\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}), \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega - (\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega. \end{aligned} \quad (4.23)$$

The last term above can be rewritten as

$$\begin{aligned} (\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega &= (\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega - (\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega \\ &= (\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})_\Omega - (\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega, \end{aligned}$$

where we used (4.21c). The penultimate term above can be estimated using Lemma 4.4 and triangle inequality as follows.

$$\begin{aligned} (\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})_\Omega &\leq C (\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}) \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} \\ &\leq C \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)}^2 + C (\|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}) \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} \end{aligned}$$

With these observations, we return to (4.23). Using the positive definiteness of  $\mathcal{A}$  also, we find that

$$\begin{aligned} \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)}^2 &\leq C (\|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}) \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} \\ &\quad + C \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)} \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} + C \|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)}^2. \end{aligned}$$

By standard manipulations using an arithmetic-geometric mean inequality, we can (kick-back and) remove the term  $\|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)}$  from the right hand side. Thus,

$$\|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} \leq C (\|\underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}).$$

The estimate for  $\underline{\boldsymbol{\sigma}}^h$  follows by applying the triangle inequality.

To prove the estimate for  $\underline{\boldsymbol{\rho}}^h$ , we now only need to apply Lemma 4.4 and the already proved estimate for  $\underline{\boldsymbol{\sigma}}^h$ .  $\square$

Now we will prove an error estimate for  $\mathbf{u}$  by a duality argument. For this, we need the following dual problem:

$$\operatorname{div} \underline{\boldsymbol{\psi}} = \boldsymbol{\theta} \quad \text{in } \Omega, \quad (4.24a)$$

$$\mathcal{A} \underline{\boldsymbol{\psi}} - \boldsymbol{\epsilon}(\boldsymbol{\phi}) = 0 \quad \text{in } \Omega, \quad (4.24b)$$

$$\boldsymbol{\phi} = 0 \quad \text{on } \partial\Omega. \quad (4.24c)$$

We assume that  $\underline{\boldsymbol{\psi}} \in \operatorname{dom} \underline{\mathbf{\Pi}}$ . We also assume that

$$\|\underline{\boldsymbol{\psi}}\|_{H^s(\Omega)} + \|\boldsymbol{\phi}\|_{H^{1+s}(\Omega)} \leq C \|\boldsymbol{\theta}\|_{L^2(\Omega)}, \quad (4.25)$$

for some  $0 \leq s \leq 1$ . Such inequalities are known to hold in several cases, e.g., in the case of planar elasticity with scalar coefficients on a convex domain, it holds [8] with  $s = 1$ .

**Theorem 4.5.** *If the regularity estimate (4.25) holds, then we have*

$$\|\mathbf{P} \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \leq C h^r (\|\underline{\boldsymbol{\sigma}} - \underline{\mathbf{\Pi}} \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} + \|\underline{\boldsymbol{\rho}} - \underline{\mathbf{P}} \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}),$$

where  $h = \max(h_K : K \in \Omega_h)$  and

$$r = \begin{cases} s, & \text{if } k \geq 2, \\ 0, & \text{if } k = 1. \end{cases}$$

*Proof.* This proof is substantially similar to a proof in [15], so we will only mention the differences.

First, consider the dual problem (4.24) with  $\boldsymbol{\theta} = \mathbf{P}\mathbf{u} - \mathbf{u}^h$ . Then, since  $\mathbf{P}\operatorname{div} \underline{\boldsymbol{\psi}} = \operatorname{div} \underline{\boldsymbol{\psi}}$ ,

$$(\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\eta}})_K = (\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\eta}})_K = 0, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\mathbf{A}}^k(K),$$

by Theorem 3.1(2). With this observation, proceeding as in the proof of [15, Lemma 5.2], we can derive the identity

$$\begin{aligned} (\mathbf{P}\mathbf{u} - \mathbf{u}^h, \boldsymbol{\theta})_\Omega &= (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\psi} - \underline{\boldsymbol{\Pi}}\boldsymbol{\psi})_\Omega + (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \underline{\boldsymbol{\xi}} - \underline{\mathbf{P}}\underline{\boldsymbol{\xi}})_\Omega \\ &\quad - (\boldsymbol{\sigma} - \underline{\boldsymbol{\Pi}}\boldsymbol{\sigma}, \operatorname{grad}(\boldsymbol{\phi} - \mathbf{P}\boldsymbol{\phi}))_\Omega + (\underline{\boldsymbol{\rho}} - \underline{\mathbf{P}}\underline{\boldsymbol{\rho}}, \underline{\boldsymbol{\Pi}}\boldsymbol{\psi} - \boldsymbol{\psi})_\Omega. \end{aligned}$$

where the gradient is taken element by element. Note that in the lowest order case,  $\mathbf{P}\boldsymbol{\phi}$  is piecewise constant (so  $\operatorname{grad}(\boldsymbol{\phi} - \mathbf{P}\boldsymbol{\phi})$  provides no extra power of  $h$ ). With this in mind, the stated estimate follows using standard approximation properties for the  $L^2$ -projections and Theorem 3.1(3) for  $\underline{\boldsymbol{\Pi}}$ .  $\square$

## 5. CONNECTION WITH STOKES ELEMENTS

It is no secret that there is a close connection between elasticity elements and Stokes elements [2, 12, 17] in two dimensions. For example, the relationship between the PEERS elasticity element and the Stokes MINI element was clarified in [17] and again, more recently, in [12]. In [12] such relationships were studied further leading to explicit conditions on properties of Fortin-like projectors. Our purpose in this section is to briefly recall this connection and to put our contribution in two dimensions in perspective. More importantly, this section will set the stage for the following section where we show that some weakly imposed symmetry methods in fact produce stress tensors that are strongly symmetric. The results in this section are essentially contained in [12, 16, 17].

First, consider the two-dimensional (2D) case. Suppose we want to construct a stress finite element space  $\widehat{\mathbf{V}}^h$  (keeping  $\underline{\mathbf{A}}^h$  and  $\mathbf{W}^h$  as defined previously) using a Stokes pair of spaces  $\mathbf{S}^h \times R^h \subseteq \mathbf{H}^1(\Omega) \times L^2(\Omega)$  (velocity and pressure spaces, respectively). We set the pressure space  $R^h$  to  $\{r : r|_K \in \mathcal{P}^k(K) \text{ for all elements } K\}$ , so that it can put into an isomorphism with  $\underline{\mathbf{A}}^h$  by the map  $\underline{\mathbf{X}}_2 : R^h \mapsto \underline{\mathbf{A}}^h$  defined by

$$\underline{\mathbf{X}}_2 w = w \underline{\boldsymbol{\chi}}, \quad \text{where } \underline{\boldsymbol{\chi}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then it is easy to see that the following diagram commutes (an easy consequence of a commuting diagram in [16]):

$$\begin{array}{ccccc} R^h & \xrightarrow{\frac{1}{2}\underline{\mathbf{X}}_2} & \underline{\mathbf{A}}^h & & \\ P \operatorname{div} \uparrow & & \uparrow \underline{\mathbf{P}} \operatorname{skw} & & \\ \mathbf{S}^h & \xrightarrow{\operatorname{curl}} & \widehat{\mathbf{V}}^h & \xrightarrow{\operatorname{div}} & \mathbf{W}^h \longrightarrow 0. \end{array} \quad (5.26)$$

Here  $\operatorname{skw}(\cdot)$  denotes the operator mapping matrices to their skew-symmetric parts, and  $P$  denotes the  $L^2$ -orthogonal projection onto  $R^h$  (and  $\underline{\mathbf{P}}$ , as before, denotes the  $L^2$ -orthogonal projection onto  $\underline{\mathbf{A}}^h$ ). The commutativity in the diagram (5.26) can be alternately expressed as

$$\frac{1}{2}\underline{\mathbf{X}}_2 P \operatorname{div} \mathbf{s}^h = \underline{\mathbf{P}} \operatorname{skw} \operatorname{curl} \mathbf{s}^h \quad (5.27)$$

for all  $\mathbf{s}^h$  in  $\mathbf{S}^h$ . Note that it is implicit in the diagram that we require  $\text{curl } \mathbf{S}^h \subseteq \widehat{\mathbf{V}}^h$  and the map  $\text{div}(\cdot)$  to be a surjection from  $\widehat{\mathbf{V}}^h$  onto  $\mathbf{W}^h$ .

**Proposition 5.1.** *Assume the following:*

(1)  $\widehat{\mathbf{V}}^h$  is such that

$$\text{curl } \mathbf{S}^h \subseteq \widehat{\mathbf{V}}^h. \quad (5.28)$$

(2)  $\mathbf{H}^1$ -continuous right inverse of  $\text{div}$  in  $R^h$ : For any  $r^h$  in  $R^h$  there is a  $\mathbf{s}^h$  in  $\mathbf{S}^h$  such that

$$P \text{div } \mathbf{s}^h = r^h, \quad \|\mathbf{s}^h\|_{\mathbf{H}^1} \leq C \|r^h\|_{L^2}. \quad (5.29)$$

(3)  $\mathbf{L}^2$ -continuous right inverse of  $\text{div}$  in  $\mathbf{W}^h$ : For any  $\mathbf{w}^h$  in  $\mathbf{W}^h$ , there is a  $\underline{\boldsymbol{\tau}}^h$  in  $\widehat{\mathbf{V}}^h$  such that

$$\text{div } \underline{\boldsymbol{\tau}}^h = \mathbf{w}^h, \quad \|\underline{\boldsymbol{\tau}}^h\|_{H(\text{div})} \leq C \|\mathbf{w}^h\|_{L^2}. \quad (5.30)$$

Then, for any  $\mathbf{w} \in \mathbf{W}^h$  and  $\boldsymbol{\zeta} \in \mathbf{A}^h$ , there exists a  $\underline{\boldsymbol{\tau}}$  in  $\widehat{\mathbf{V}}^h$  satisfying (4.19), i.e., the spaces  $\widehat{\mathbf{V}}^h, \mathbf{W}^h, \mathbf{A}^h$  satisfy the elasticity inf-sup condition.

*Proof.* We find the required  $\underline{\boldsymbol{\tau}}$  in two steps. First, by the given condition (5.30), we find  $\underline{\boldsymbol{\tau}}_1$  in  $\widehat{\mathbf{V}}^h$  such that

$$\text{div } \underline{\boldsymbol{\tau}}_1 = \mathbf{w}^h, \quad \|\underline{\boldsymbol{\tau}}_1\|_{H(\text{div})} \leq C \|\mathbf{w}^h\|_{L^2}. \quad (5.31)$$

Second, by (5.29), we find  $\mathbf{s}^h$  in  $\mathbf{S}^h$  such that

$$P \text{div } \mathbf{s}^h = 2\underline{\mathbf{X}}_2^{-1} \underline{\mathbf{P}}(\boldsymbol{\zeta} - \text{skw } \underline{\boldsymbol{\tau}}_1), \quad \|\mathbf{s}^h\|_{\mathbf{H}^1} \leq C \|\boldsymbol{\zeta} - \text{skw } \underline{\boldsymbol{\tau}}_1\|_{L^2}.$$

Setting  $\underline{\boldsymbol{\tau}}_2 = \text{curl } \mathbf{s}^h$  and using (5.26), this implies

$$\underline{\mathbf{P}} \text{skw } \underline{\boldsymbol{\tau}}_2 = \underline{\mathbf{P}}(\boldsymbol{\zeta} - \text{skw } \underline{\boldsymbol{\tau}}_1), \quad \|\underline{\boldsymbol{\tau}}_2\|_{H(\text{div})} \leq C(\|\boldsymbol{\zeta}\|_{L^2} + \|\underline{\boldsymbol{\tau}}_1\|_{L^2}) \quad (5.32)$$

Setting  $\underline{\boldsymbol{\tau}} = \underline{\boldsymbol{\tau}}_1 + \underline{\boldsymbol{\tau}}_2$  and combining (5.31) and (5.32), we prove the result.  $\square$

*Example 5.2.* In [17, Example 4.2], we find that the lowest order case of our 2D element is already discovered by considering the conforming “ $P_2^+ P_{-1}$ ”-Crouzeix-Raviart element. To explain this, first set the Stokes pair to the Crouzeix-Raviart element by

$$\begin{aligned} R^h &= \{r \in L^2(\Omega) : r|_K \in \mathcal{P}^1(K), \text{ for all elements } K\}, \\ \mathbf{S}^h &= \{\mathbf{s}^h \in \mathbf{H}^1(\Omega) : \mathbf{s}^h|_K \in \mathcal{P}^2(K) + b_K\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right) \text{ for all elements } K\} \end{aligned}$$

for some constants  $\alpha$  and  $\beta$ . To use Proposition 5.1 so as to obtain a stress element  $\widehat{\mathbf{V}}^h$ , we must satisfy (5.28), so it is natural to set

$$\widehat{\mathbf{V}}(K) = \underline{\mathcal{P}}^1(K) + \text{curl}\left(b_K\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right)\right)$$

and  $\widehat{\mathbf{V}}^h = \{\mathbf{v} \in \mathbf{H}(\text{div}, \Omega) : \mathbf{v}|_K \in \widehat{\mathbf{V}}(K), \text{ for all } K \in \Omega_h\}$ . This is indeed the same space we introduced previously, i.e.,  $\widehat{\mathbf{V}}^h = \mathbf{V}^h$  when  $k = 1$  as

$$\{\text{curl}\left(b_K\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right)\right) : \alpha, \beta \in \mathbb{R}\} = \{\text{curl}\left(\text{curl}\left(\begin{smallmatrix} 0 & -\gamma \\ \gamma & 0 \end{smallmatrix}\right)b_K\right) : \gamma \in \mathcal{P}^1(K)\}.$$

Finally note that the closely related Amara-Thomas elements [2] are the same as the our 2D elements in [15]. In the lowest order case, they take the form  $\widehat{\mathbf{V}}(K) = \underline{\mathbf{RT}}^1(K) + \text{curl}\left(b_K\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right)\right)$ .  $\square$

Next, let us consider the three-dimensional (3D) case. Now, we must put together three copies of Stokes spaces to form matrix Stokes elements. Namely, let  $\underline{\mathbf{S}}^h$  denote an  $\underline{\mathbf{H}}^1(\Omega)$ -conforming space of matrix functions and  $\mathbf{R}^h \subseteq \mathbf{L}^2(\Omega)$ . To obtain an analogue of (5.26), we first set  $\mathbf{R}^h$  to a space homeomorphic to  $\underline{\mathbf{A}}^h$  using the map  $\underline{\mathbf{X}}_3$  (in place of  $\underline{\mathbf{X}}_2$  in (5.26)) defined by

$$\underline{\mathbf{X}}_3 \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix},$$

i.e., we set  $\mathbf{R}^h = \underline{\mathbf{X}}_3^{-1} \underline{\mathbf{A}}^h$ . Then, letting  $\text{tr}(\cdot)$  denote the trace, and defining

$$\underline{\Xi} \underline{\mathbf{v}} = \underline{\mathbf{v}}^t - (\text{tr } \underline{\mathbf{v}}) \underline{\mathbf{I}}, \quad \underline{\Xi}^{-1} = \underline{\mathbf{v}}^t - \frac{1}{2} (\text{tr } \underline{\mathbf{v}}) \underline{\mathbf{I}},$$

we recall the well known 3D analogue of (5.27), namely

$$\text{skw curl } \underline{\mathbf{w}} = \frac{1}{2} \underline{\mathbf{X}}_3 \text{div } \underline{\Xi} \underline{\mathbf{w}}.$$

Set  $\widehat{\underline{\mathbf{S}}}^h = \underline{\Xi}^{-1} \underline{\mathbf{S}}^h$ . Introducing appropriate projections, it easily follows that the diagram

$$\begin{array}{ccccc} & & \mathbf{R}^h & \xrightarrow{\frac{1}{2} \underline{\mathbf{X}}_3} & \underline{\mathbf{A}}^h \\ & \uparrow \mathbf{P} \text{ div} & & & \uparrow \mathbf{P} \text{ skw} \\ & \underline{\mathbf{S}}^h & & & \\ & \uparrow \underline{\Xi} & & & \\ \widehat{\underline{\mathbf{S}}}^h & \xrightarrow{\text{curl}} & \widehat{\underline{\mathbf{V}}}^h & \xrightarrow{\text{div}} & \mathbf{W}^h \longrightarrow 0 \end{array}$$

commutes. Note that the bottom row is not required to be an exact sequence. (Indeed, typical exact sequences have a curl-conforming space in place of the  $\underline{\mathbf{H}}^1$ -conforming  $\widehat{\underline{\mathbf{S}}}^h$ .) We only need  $\text{curl } \widehat{\underline{\mathbf{S}}}^h \subseteq \widehat{\underline{\mathbf{V}}}^h$ . With the above diagram, the following result is proved along the same lines of the proof of Proposition 5.1 (so we omit the details).

**Proposition 5.3.** *The statement of Proposition 5.1 holds in three dimensions after replacing (5.28) by*

$$\text{curl } \widehat{\underline{\mathbf{S}}}^h \subseteq \widehat{\underline{\mathbf{V}}}^h, \tag{5.33}$$

and replacing  $P, R^h$ , and  $\mathbf{S}^h$  by  $\mathbf{P}, \mathbf{R}^h$ , and  $\underline{\mathbf{S}}^h$ , respectively.

## 6. SERENDIPITOUS EXACT SYMMETRY

By way of a final example using the Stokes connection, we now show that in certain type of meshes, we may omit the bubbles  $B(\underline{\mathbf{A}}^k)$  and still obtain a stable method. This is a particularly interesting example, as the removal of bubbles implies that the test and trial space in the weak symmetry constraint (1.5c) use polynomials of the same degree. Therefore, the method yields stress approximations that are *exactly* symmetric. It is well known that stress elements with exact symmetry do not come easy. Known elements [1, 3, 7] have too many degrees of freedom. So any other game yielding exactly symmetric stress approximations is worth a pursuit.

First we consider the 2D case. We recall the conditions on planar meshes that were discovered in [23, 25] in the context of exactly incompressible Stokes flow approximations in the assumption below. Its relevance in elasticity will be clear shortly.

*Assumption 6.1* (Away from singular vertices [23, 25]). We assume that the planar meshes are quasiuniform and satisfy the following two conditions:

- (1) No interior mesh vertex is “singular” in the sense of [25]. An interior vertex is called singular if all the edges meeting at this vertex fall on two straight lines.
- (2) No vertex is close to singularity in the following sense [23]: Let  $\theta_i$ ,  $i = 1, \dots, m$ , be the angles of the triangles meeting at one vertex  $x$ . Define  $R(x) = \max\{|\theta_i + \theta_j - \pi| : 1 \leq i, j \leq m, i - j = 1 \pmod{m}\}$  for interior mesh vertices  $x$ . For boundary vertices, the same definition holds verbatim after “mod” is deleted. We assume that there is a fixed  $\delta > 0$  independent of  $h$  such that

$$R(x) \geq \delta > 0$$

for all vertices  $x$  of all meshes under consideration.

- (3) The polynomial degree satisfies  $k \geq 3$ .

One class of meshes for which this assumption holds is the so called Hsieh-Clough-Toucher grids, as observed in [6]. In fact, they also showed that for such grids, the associated polynomial degree can be further reduced, as stated in the next assumption below.

A complete generalization of the above mentioned concept of singular mesh objects for 3D tetrahedral meshes remains unknown. However in [26], the 3D version of the Hsieh-Clough-Toucher grids is shown to provide a class of meshes without singular vertices and edges, leading us to consider 3D meshes also in the following assumption:

*Assumption 6.2* (Hsieh-Clough-Toucher grids [6, 26]). In either the 2D (triangular) or the 3D (tetrahedral) case, assume the following:

- (1) The mesh is obtained from a quasiuniform mesh after splitting each of its elements into  $d + 1$  elements by connecting the vertices of the element to its barycenter.
- (2) In the 2D case, assume  $k \geq 1$ .
- (3) In the 3D case, assume  $k \geq 2$ .

**Theorem 6.3.** *Suppose either Assumption 6.1 or Assumption 6.2 holds. Consider the mixed method (1.5) with the local spaces*

$$\underline{\mathbf{V}}(K) = \underline{\mathcal{P}}^k(K), \quad \mathbf{W}(K) = \mathcal{P}^{k-1}(K), \quad \underline{\mathbf{A}}(K) = \underline{\mathcal{A}}^k(K),$$

(and the degree  $k$  as specified in each case in the assumptions). The method is uniquely solvable, is stable, and quasioptimal. Moreover, its stress approximation  $\underline{\boldsymbol{\sigma}}^h$  is exactly symmetric.

*Proof.* Consider the 2D case first. We apply Proposition 5.1 after setting the Stokes pair to

$$\begin{aligned} R^h &= \{r \in L^2(\Omega) : r|_K \in \mathcal{P}^k(K), \text{ for all elements } K\}, \\ \mathbf{S}^h &= \{\mathbf{s} \in \mathbf{H}^1(\Omega) : \mathbf{s}|_K \in \mathcal{P}^{k+1}(K) \text{ for all elements } K\}. \end{aligned}$$

Then, if Assumption 6.1 holds, by [23, Theorem 5.1], there is an  $\mathbf{H}^1$ -continuous right inverse for the divergence operator for  $k \geq 3$ . If Assumption 6.2 holds, then the same

result follows from [6] for  $k \geq 1$ . Hence, the choice

$$\underline{\mathbf{V}}^h = \{\underline{\boldsymbol{\tau}} \in H(\operatorname{div}) : \underline{\boldsymbol{\tau}} \in \underline{\mathcal{P}}^k(K) \text{ for all elements } K\}$$

satisfies all the three conditions of Proposition 5.1. Consequently, the method is stable and Babuška-Brezzi theory gives quasioptimality. The stated exact symmetry of  $\underline{\boldsymbol{\sigma}}^h$  follows because  $\underline{\mathbf{V}}(K)$  and  $\underline{\mathbf{A}}(K)$  consist of polynomials of the same degree.

In the 3D case of Assumption 6.2, we apply Proposition 5.3 instead, with

$$\underline{\mathbf{S}}^h = \{\underline{\mathbf{s}} \in \underline{\mathbf{H}}^1(\Omega) : \underline{\mathbf{s}}|_K \in \underline{\mathcal{P}}^{k+1}(K) \text{ for all elements } K\}.$$

Note that  $\widehat{\underline{\mathbf{S}}}^h$  and  $\underline{\mathbf{S}}^h$  coincide for the above choice of  $\underline{\mathbf{S}}^h$ , so (5.33) holds. Moreover, the analysis in [26] shows that for any  $\mathbf{r}^h$  in  $\mathbf{R}^h \equiv \underline{\mathbf{X}}_3^{-1} \underline{\mathbf{A}}_h$ , there is a  $\underline{\mathbf{s}}^h$  in  $\underline{\mathbf{S}}^h$  satisfying  $\mathbf{P} \operatorname{div} \underline{\mathbf{s}}^h = \mathbf{r}^h$ , and  $\|\underline{\mathbf{s}}^h\|_{H^1} \leq C \|\mathbf{r}^h\|_{L^2}$ . (We note that although the result in [26] is stated only for spaces with zero boundary condition, the analysis there holds more generally even for the space  $\underline{\mathbf{S}}^h$  without the boundary condition.) Thus, the conditions of Proposition 5.3 can be verified and the proof can be finished as in the 2D case.  $\square$

With a more refined analysis, we can separate the errors in the displacement from that of the stress and consequently prove optimal estimates for the latter, as we show next.

**Theorem 6.4.** *With the same hypothesis as the previous theorem we have*

$$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)} \leq C \left( \|\underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)} \right).$$

*Proof.* In the proof of Theorem 6.3, we showed that the inf-sup condition of Lemma 4.1 holds under the given assumptions. Therefore, we can follow along the lines of the proof of Lemma 4.4 to get

$$\|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)} \leq C \left( \|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega)} + \|\underline{\mathbf{P}} \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)} \right). \quad (6.34)$$

Next, to control  $\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$ , we start by observing that from the second error equation (4.21b) and the properties of  $\underline{\mathbf{I}}_{\text{BDM}}$  (see (3.16) for definition) that

$$\operatorname{div} (\underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h) = 0.$$

Therefore, using (4.21a) with  $\underline{\mathbf{v}} = \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$  (note that  $\underline{\mathbf{v}}$  does in fact belong to  $\underline{\mathbf{V}}^h$ ) we get

$$(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h), \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega = -(\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega$$

By Theorem 6.3,  $\underline{\boldsymbol{\sigma}}^h$  is symmetric, so the last term can be rewritten as

$$(\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h), \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega = -(\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})_\Omega.$$

By the positive-definiteness of  $\mathcal{A}$ , we thus have

$$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{\mathcal{A}}^2 = (\mathcal{A}(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h), \underline{\boldsymbol{\sigma}} - \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}})_\Omega - (\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\mathbf{I}}_{\text{BDM}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}})_\Omega$$

where  $\|\underline{\mathbf{v}}\|_{\mathcal{A}}^2 = (\mathcal{A}\underline{\mathbf{v}}, \underline{\mathbf{v}})_\Omega$ . Now, estimating by elementary inequalities like the Cauchy-Schwarz inequality, and using (6.34), we finish the proof.  $\square$



## 7. SOME IMPLEMENTATION ASPECTS

We conclude by briefly discussing dimension-reduced implementation by hybridization and an accuracy-enhancing postprocessing technique.

A standard implementation of the mixed method (1.5) results in a large indefinite matrix system (for the coefficients in a basis) for  $\underline{\boldsymbol{\sigma}}^h$ ,  $\mathbf{u}^h$ , and  $\underline{\boldsymbol{\rho}}^h$ . Hybridization can reduce the size of this system, as well as provide a symmetric positive definite reduced system. To discuss this, we start, perhaps counterintuitively, by enlarging the mixed method (1.5) to include an additional unknown  $\boldsymbol{\lambda}^h$  in

$$\mathbf{M}^h = \{\boldsymbol{\mu} : \boldsymbol{\mu}|_F \in \mathcal{P}^k(F) \text{ for all mesh faces } F \in \mathcal{E}_h, \text{ and } \boldsymbol{\mu}|_{\partial\Omega} = 0\},$$

and enlarging the space  $\underline{\mathbf{V}}^h$  by removing its  $\underline{\mathbf{H}}(\text{div}, \Omega)$ -continuity constraints to obtain the space

$$\tilde{\underline{\mathbf{V}}}^h = \{\underline{\mathbf{v}} : \underline{\mathbf{v}}|_K \in \underline{\mathbf{V}}(K) \text{ for all mesh elements } K \in \Omega_h\}.$$

Note that for all the choices of  $\underline{\mathbf{V}}(K)$  we considered in the previous sections, we have

$$\begin{aligned} \underline{\mathbf{V}}(K) &\subset \underline{\mathcal{P}}^k(K) \\ \underline{\mathbf{v}} \mathbf{n}|_F &\in \mathcal{P}^k(F) \quad \text{for all faces } F \text{ of } K \text{ and for all } \underline{\mathbf{v}} \in \underline{\mathbf{V}}(K), \end{aligned}$$

hence  $\underline{\mathbf{V}}^h = \{\underline{\mathbf{v}} \in \tilde{\underline{\mathbf{V}}}^h : \langle \underline{\mathbf{v}} \mathbf{n}, \boldsymbol{\mu} \rangle_{\partial\Omega_h} = 0 \text{ for all } \boldsymbol{\mu} \in \mathbf{M}^h\}$ , where

$$\langle \omega, v \rangle_{\partial\Omega_h} := \sum_{K \in \Omega_h} \langle \omega, v \rangle_{\partial K}.$$

Thus the space  $\mathbf{M}^h$  is the space of Lagrange multipliers corresponding to the  $\underline{\mathbf{H}}(\text{div}, \Omega)$ -continuity constraints of  $\underline{\mathbf{V}}^h$ .

Hybridization removes the stress interelement continuity constraints from the global finite space and reimposes them as an equation of the method, i.e., the solution of hybridized method,  $(\underline{\boldsymbol{\sigma}}^h, \mathbf{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h) \in \tilde{\underline{\mathbf{V}}}^h \times \mathbf{W}^h \times \underline{\mathbf{A}}^h \times \mathbf{M}^h$ , satisfies

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}^h, \underline{\mathbf{v}})_{\Omega_h} + (\mathbf{u}^h, \text{div } \underline{\mathbf{v}})_{\Omega_h} + (\underline{\boldsymbol{\rho}}^h, \underline{\mathbf{v}})_{\Omega_h} + \langle \boldsymbol{\lambda}^h, \underline{\mathbf{v}} \mathbf{n} \rangle_{\partial\Omega_h} = 0, \quad (7.35a)$$

$$(\text{div } \underline{\boldsymbol{\sigma}}^h, \boldsymbol{\omega})_{\Omega_h} = (\mathbf{f}, \boldsymbol{\omega})_{\Omega_h}, \quad (7.35b)$$

$$(\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\eta}})_{\Omega_h} = 0, \quad (7.35c)$$

$$\langle \underline{\boldsymbol{\sigma}}^h \mathbf{n}, \boldsymbol{\mu} \rangle_{\partial\Omega_h} = 0, \quad (7.35d)$$

for all  $(\underline{\mathbf{v}}, \boldsymbol{\omega}, \underline{\boldsymbol{\eta}}, \boldsymbol{\mu}) \in \tilde{\underline{\mathbf{V}}}^h \times \mathbf{W}^h \times \underline{\mathbf{A}}^h \times \mathbf{M}^h$ . Here, differential operators are applied element by element and

$$(\omega, v)_{\Omega_h} := \sum_{K \in \Omega_h} (\omega, v)_K.$$

The following result can be proved as in [15].

**Proposition 7.1.** *There is a unique  $(\underline{\boldsymbol{\sigma}}^h, \mathbf{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h) \in \tilde{\underline{\mathbf{V}}}^h \times \mathbf{W}^h \times \underline{\mathbf{A}}^h \times \mathbf{M}^h$  satisfying (7.35). Moreover, the first three components of the solution, namely  $(\underline{\boldsymbol{\sigma}}^h, \mathbf{u}^h, \underline{\boldsymbol{\rho}}^h)$  coincide with that of the mixed method (1.5).*

The purpose of discussing the enlarged system (7.35) is that it allows elimination of all the unknowns (for all choices of stress spaces we considered) except  $\boldsymbol{\lambda}^h$ . Specifically, there is a locally computable symmetric coercive bilinear form  $a_h(\cdot, \cdot)$  and a linear form  $\ell_h(\cdot)$  on  $\mathbf{M}^h$  such that  $\boldsymbol{\lambda}^h$  is the unique solution of

$$a_h(\boldsymbol{\lambda}^h, \boldsymbol{\mu}) = \ell(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \mathbf{M}^h. \quad (7.36)$$

This statement can be proved and the forms explicitly written down by the techniques in [14] (see also [15]). Thus the preferred method of implementation is to solve (7.36) (a sparse symmetric positive definite system) first for  $\boldsymbol{\lambda}^h$ . The other unknowns  $\underline{\boldsymbol{\sigma}}^h$ ,  $\mathbf{u}^h$  and  $\underline{\boldsymbol{\rho}}^h$  can then be locally computed using the  $\boldsymbol{\lambda}^h$  as shown in [14, 15].

Finally, we adapt Stenberg's postprocessing [24] to our methods, to obtain a better displacement approximation. On each simplex  $K \in \Omega_h$ , the post-processed approximation  $\mathbf{u}^{h,*}$  is the (unique) function in  $\mathcal{P}^{k+1}(K)$  satisfying

$$\begin{aligned} (\text{grad } \mathbf{u}^{h,*}, \text{grad } \boldsymbol{\omega})_K &= (\mathcal{A}\underline{\boldsymbol{\sigma}}^h + \underline{\boldsymbol{\rho}}^h, \text{grad } \boldsymbol{\omega})_K \quad \forall \boldsymbol{\omega} \in \mathcal{P}_\perp^{k+1, k-1}(K), \\ (\mathbf{u}^{h,*}, \mathbf{w})_K &= (\mathbf{u}^h, \mathbf{w})_K \quad \forall \mathbf{w} \in \mathcal{P}^{k-1}(K), \end{aligned}$$

where  $\mathcal{P}_\perp^{k+1, k-1}(K) := \{\mathbf{v} \in \mathcal{P}^{k+1}(K) : (\mathbf{v}, \boldsymbol{\omega})_K = 0 \text{ for all } \boldsymbol{\omega} \in \mathcal{P}^{k-1}(K)\}$ . Then, taking advantage of Theorem 4.5 (we note that this result also holds for the methods considered in Section 6), we can prove that the local post-processing scheme gives an improved approximation for  $\mathbf{u}$ .

**Theorem 7.2.** *Suppose  $\mathbf{u}$  is in  $\mathbf{H}^{k+2}(\Omega)$  and the regularity assumption (4.25) holds with  $s = 1$ . Then*

$$\|\mathbf{u} - \mathbf{u}^{h,*}\|_{L^2(\Omega)} \leq C h^{k+\ell} \|\mathbf{u}\|_{\mathbf{H}^{k+2}(\Omega)},$$

where  $\ell = 1$  if  $k \geq 2$  and  $\ell = 0$  if  $k = 1$ .

The proof is similar to the analogous result in Stenberg [24] hence we omit the details.

#### ACKNOWLEDGEMENTS

The results of Section 6 are due to the authors and Professor Richard Falk. The authors appreciate his encouragement and the opportunities he offered for discussing these results.

#### 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 AND R. WINThER, *Finite elements for symmetric tensors in three dimensions*, Math. Comp. 77 (2008), no. 263, 1229–1251.
- [4] 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.
- [5] 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.
- [6] D. N. ARNOLD AND J. QIN, *Quadratic velocity/linear pressure Stokes elements*, in Advances in Computer Methods for Partial Differential Equations-VII, R. Vichnevetsky, D. Knight, and G. Richter, eds., IMACS, 1992, pp. 28–34.
- [7] D. N. ARNOLD AND R. WINThER, *Mixed finite elements for elasticity*, Numer. Math., 92 (2002), pp. 401–419.

- [8] 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.
- [9] F. BREZZI, J. DOUGLAS AND L.D. MARINI, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. 47 (1985), no. 2, 217–235.
- [10] F. BREZZI, J. DOUGLAS, R. DURÁN AND M. FORTIN, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math. 51 (1987), no. 2, 237–250.
- [11] F. BREZZI AND M. FORTIN, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.
- [12] D. BOFFI, F. BREZZI AND M. FORTIN, *Reduced symmetry elements in linear elasticity*, Commun. Pure Appl. Anal. 8 (2009), no. 1, 95–121.
- [13] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, Amsterdam, 1978.
- [14] B. COCKBURN AND J. GOPALAKRISHNAN, *A characterization of hybridized mixed methods for the Dirichlet problem*, SIAM J. Numer. Anal., 42 (2004), pp. 283–301.
- [15] B. COCKBURN, J. GOPALAKRISHNAN, AND J. GUZMÁN, *A new elasticity element made for enforcing weak stress symmetry*, Math. Comp., 79 (2010), 1331–1349.
- [16] R. FALK, *Finite elements for linear elasticity*, in Mixed Finite Elements: Compatibility Conditions (eds. D. Boffi and L. Gastaldi), Lecture Notes in Math., 1939 Springer-Verlag, Heidelberg, 2008.
- [17] 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.
- [18] J. GOPALAKRISHNAN, L. E. GARCÍA-CASTILLO, AND L. F. DEMKOWICZ, *Nédélec spaces in affine coordinates*, Comput. Math. Appl., 49 (2005), pp. 1285–1294.
- [19] M. MORLEY, *A family of mixed finite elements for linear elasticity* Numer. Math. 55 (1989), no. 6, 633–666.
- [20] J.-C. NÉDÉLEC, *Mixed Finite Elements in  $\mathbb{R}^3$* , Numer. Math., 35 (1980), pp. 315–341.
- [21] J.-C. NÉDÉLEC, *A new family of mixed finite elements in  $\mathbb{R}^3$*  Numer. Math. 50 (1986), no. 1, 57–81.
- [22] P.-A. RAVIART AND J. M. THOMAS, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), Springer, Berlin, 1977, pp. 292–315. Lecture Notes in Math., Vol. 606.
- [23] L. R. SCOTT AND M. VOGELIUS, *Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 111–143.
- [24] R. STENBERG, *A family of mixed finite elements for the elasticity problem*, Numer. Math. 53 (1988), no. 5, 513–538.
- [25] M. VOGELIUS, *A right-inverse for the divergence operator in spaces of piecewise polynomials. Application to the p-version of the finite element method*, Numer. Math., 41 (1983), pp. 19–37.
- [26] S. ZHANG, *A new family of stable mixed finite elements for the 3D Stokes equations*, Math. Comp., 74 (2005), pp. 543–554 (electronic).

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