2009

A New Elasticity Element Made for Enforcing Weak Stress Symmetry

Bernardo Cockburn
University of Minnesota - Twin Cities

Jay Gopalakrishnan
Portland State University, gjay@pdx.edu

Johnny Guzmán
Brown University

A NEW ELASTICITY ELEMENT MADE FOR ENFORCING WEAK STRESS SYMMETRY

BERNARDO COCKBURN, JAYADEEP GOPALAKRISHNAN, AND JOHNNY GUZMÁN

ABSTRACT. We introduce a new mixed method for linear elasticity. The novelty is a simplicial element for the approximate stress. For every positive integer k, the row-wise divergence of the element space spans the set of polynomials of total degree k. The degrees of freedom are suited to achieve continuity of the normal stresses. What makes the element distinctive is that its dimension is the smallest required for enforcing a weak symmetry condition on the approximate stress. This is achieved using certain "bubble matrices", which are special divergence-free matrix-valued polynomials. We prove that the approximation error is of order $k + 1$ in both the displacement and the stress, and that a postprocessed displacement approximation converging at order $k + 2$ can be computed element by element. We also show that the globally coupled degrees of freedom can be reduced by hybridization to those of a displacement approximation on the element boundaries.

1. INTRODUCTION

In this paper, we introduce a new mixed method for the system of equations describing linear elasticity, namely

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

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

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

where $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) is a polyhedral domain. Here $\boldsymbol{u}$ is the displacement, $\underline{\boldsymbol{\sigma}}$ denotes the stress, $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) = (\operatorname{grad}\boldsymbol{u} + (\operatorname{grad}\boldsymbol{u})^t)/2$ is the symmetric part of $\operatorname{grad}\boldsymbol{u}$ and $\mathcal{A}$ is a symmetric and positive-definite tensor over the space of symmetric matrices. Differential operators are applied row by row, i.e., in (1.1a), the ith row of $\operatorname{div}\underline{\boldsymbol{\sigma}}$ is the divergence of the ith row vector of the matrix $\underline{\boldsymbol{\sigma}}$. Similarly, the ith row of the matrix $\operatorname{grad}\boldsymbol{u}$ is the gradient (written as a row) of the ith component of the vector $\boldsymbol{u}$.

1.1. **Mixed methods for linear elasticity.** To describe the new features of the method and to facilitate comparison with previously known methods, let us introduce a general method in the mixed form for our linear elasticity problem.

Mixed methods finite elements for linear elasticity generally fall into two categories, those that enforce symmetry of the stress tensor exactly [2, 3, 6], and the

---

ones that enforce it weakly [4, 5, 18, 12, 8]. Here, we consider the latter methods. They require the introduction of the new variable $\boldsymbol{\rho} = (\operatorname{grad} \boldsymbol{u} - (\operatorname{grad} \boldsymbol{u})^t)/2$ which is the antisymmetric part of $\operatorname{grad} \boldsymbol{u}$. Then, (1.1b) becomes

$$\mathcal{A}\underline{\boldsymbol{\sigma}} - \operatorname{grad} \boldsymbol{u} + \underline{\boldsymbol{\rho}} = 0 \quad \text{in } \Omega. \tag{1.2}$$

In [4, 15], the space of approximations to $\underline{\boldsymbol{\rho}}$ are continuous or weakly continuous. In contrast, here we restrict ourselves to mixed methods for which such approximations are completely discontinuous, since this permits the element-by-element elimination of the approximation to the auxiliary variable $\underline{\boldsymbol{\rho}}$.

Given a simplicial subdivision of Ω, denoted by Ω_h, the approximation to $(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}, \underline{\boldsymbol{\rho}})$, namely $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h)$, is sought in the finite dimensional space $\underline{\boldsymbol{V}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h$, where

$$\underline{\boldsymbol{V}}^h := \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{H}}(\operatorname{div}, \Omega) : \boldsymbol{v}|_K \in \underline{\boldsymbol{V}}(K), \text{ for all } K \in \Omega_h\}, \tag{1.3a}$$

$$\boldsymbol{W}^h := \{\boldsymbol{w} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{w}|_K \in \boldsymbol{W}(K), \text{ for all } K \in \Omega_h\}, \tag{1.3b}$$

$$\underline{\boldsymbol{A}}^h := \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\Omega) : \underline{\boldsymbol{v}}|_K \in \underline{\boldsymbol{A}}(K), \text{ for all } K \in \Omega_h\}. \tag{1.3c}$$

Here $\underline{\boldsymbol{L}}^2(\Omega)$ is the space of matrix-valued functions whose entries are in $L^2(\Omega)$, and $\underline{\boldsymbol{H}}(\operatorname{div}, \Omega)$ is the space of matrix-valued functions such that each row belongs to $\boldsymbol{H}(\operatorname{div}; \Omega)$ (the space of vector functions whose components and distributional divergence are in $L^2(\Omega)$), and $\boldsymbol{n}$ denotes the outward unit normal. Moreover, $\underline{\boldsymbol{V}}(K), \boldsymbol{W}(K)$ and $\underline{\boldsymbol{A}}(K)$ are finite dimensional spaces, typically composed of polynomials. The approximation is defined as the solution of

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{v}})_\Omega + (\boldsymbol{u}^h, \operatorname{div} \underline{\boldsymbol{v}})_\Omega + (\underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{v}})_\Omega = 0, \tag{1.4a}$$

$$(\operatorname{div} \underline{\boldsymbol{\sigma}}^h, \boldsymbol{\omega})_\Omega = (\boldsymbol{f}, \boldsymbol{\omega})_\Omega, \tag{1.4b}$$

$$(\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\eta}})_\Omega = 0, \tag{1.4c}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \underline{\boldsymbol{\eta}}) \in \underline{\boldsymbol{V}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h$. Above and throughout, the notation $(\boldsymbol{\zeta}, \boldsymbol{\theta})_D$ denotes the integral over D of $\boldsymbol{\zeta} \cdot \boldsymbol{\theta}$ when $\boldsymbol{\zeta}$ and $\boldsymbol{\theta}$ are vector functions, while for matrix functions $\underline{\boldsymbol{\zeta}}$ and $\underline{\boldsymbol{\theta}}$, the notation $(\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\theta}})_D$ denotes the integral over D of $\underline{\boldsymbol{\zeta}} : \underline{\boldsymbol{\theta}}$ (where "$\cdot$" and "$:$" denote the dot product and the Frobenius inner-product, respectively). As usual, for methods that enforce symmetry only weakly, we extend $\mathcal{A}$ to be a symmetric and positive definite tensor over *all* matrices, not only symmetric ones.

In the 80's, Stenberg [18] considered several methods of the form (1.4). They are extensions to the elasticity system of the Brezzi-Douglas-Marini and the Raviart-Thomas mixed methods for second-order elliptic problems. More recently, Arnold, Falk and Winther [5] also considered these methods in their exploration of how to use the de Rham complex to construct new methods for linear elasticity. Both papers set the local spaces to

$$\boldsymbol{W}(K) := \boldsymbol{W}^k(K) := \boldsymbol{\mathcal{P}}^k(K), \tag{1.5a}$$

$$\underline{\boldsymbol{A}}(K) := \underline{\boldsymbol{A}}^k(K) := \{\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{\mathcal{P}}}^k(K) : \underline{\boldsymbol{\eta}} + \underline{\boldsymbol{\eta}}^t = 0\}, \tag{1.5b}$$

where $\boldsymbol{\mathcal{P}}^k(D) := [\mathcal{P}^k(D)]^d$, $\underline{\boldsymbol{\mathcal{P}}}^k(D) := [\mathcal{P}^k(D)]^{d \times d}$ and $\mathcal{P}^k(D)$ is the space of polynomials of degree k defined on the domain D. Moreover, we use the following notation for the L^2-orthogonal complement of $\mathcal{P}^{k-1}(D)$ in $\mathcal{P}^k(D)$

$$\tilde{\mathcal{P}}^k(D) := \{v \in \mathcal{P}^k(D) : (v, w)_D = 0 \text{ for all } w \in \mathcal{P}^{k-1}(D)\},$$

and we set $\tilde{\boldsymbol{\mathcal{P}}}^k(D) := [\tilde{\mathcal{P}}^k(D)]^d$ and $\tilde{\underline{\boldsymbol{\mathcal{P}}}}^k(D) := [\tilde{\mathcal{P}}^k(D)]^{d \times d}$. For the local stress space, Stenberg's choice [18] was

$$\underline{\boldsymbol{V}}(K) := \underline{\boldsymbol{RT}}^k(K) + \delta\underline{\boldsymbol{V}}(K),$$

where $\underline{\boldsymbol{RT}}^k(K)$ is the set of matrix-valued functions on K whose rows are in the well-known [16, 17] Raviart-Thomas space $\boldsymbol{RT}^k(K) = \boldsymbol{\mathcal{P}}^k(K) + \boldsymbol{x}\mathcal{P}^k(K)$, and

$$\delta\underline{\boldsymbol{V}}(K) := \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k+d-1}(K) : \underline{\boldsymbol{v}} = \operatorname{curl}(\underline{\boldsymbol{w}}\, b_K), \underline{\boldsymbol{w}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)\}.$$

Here, curl is applied row-wise, like our other differential operators, and b_K is the *scalar bubble*, which in the three dimensional case is $b_K = \lambda_0 \lambda_1 \lambda_2 \lambda_3$ (where λ_i's denote the barycentric coordinates of the simplex K). Although Stenberg [18] only presents the Brezzi-Douglas-Marini version of his method, he argues that the Raviart-Thomas version we just described can be analyzed with his technique, and in particular, that all the approximations converge with order $k + 1$. Moreover, the approximation to $\boldsymbol{u}$ superconverges to a local projection of $\boldsymbol{u}$ which allows for a simple local post-processing procedure to compute a new approximation to $\boldsymbol{u}$ converging with order $k + 2$.

Arnold, Falk and Winther [5], in addition to (1.5), set the stress space, for all $k \geq 0$, by

$$\underline{\boldsymbol{V}}(K) := \underline{\boldsymbol{\mathcal{P}}}^{k+1}(K).$$

They showed that their approximations, like those of Stenberg, also converge with order $k + 1$. Note however that this rate is one order less than the order of the best approximation error using the above $\underline{\boldsymbol{V}}(K)$. Nonetheless, they argue that at least in the lowest-order case $k = 0$, they can reduce the space $\underline{\boldsymbol{V}}(K)$ and still obtain a stable method.

1.2. **A new element for the stress.** We now describe our new element in three dimensions ($d = 3$). Later sections will treat both the two and three-dimensional cases. We note that the two-dimensional family of elements we present here are essentially the elements developed in [1] where they use a dual-hybrid formulation. However, we give here a unified analysis of the two and three-dimensional cases which allows us to see connections. The design of our element is motivated by the following question: Can we retain the good convergence properties of Stenberg's method with a local space $\underline{\boldsymbol{V}}(K)$ of smaller dimension?

To precisely describe the issue, we consider any local stress space of the form

$$\underline{\boldsymbol{V}}(K) := \underline{\boldsymbol{RT}}^k(K) + \delta\underline{\boldsymbol{V}}(K),$$

and ask ourselves if we can find a space $\delta\underline{\boldsymbol{V}}(K)$ such that there is a projection $\underline{\boldsymbol{\Pi}}$ from the space of symmetric matrix-valued functions into $\underline{\boldsymbol{V}}_h$ satisfying the following properties:

   (i) div $\underline{\boldsymbol{\Pi}} = \boldsymbol{P}$ div ,
   (ii) $\underline{\boldsymbol{V}}(K)$ has the smallest possible dimension for all $K \in \Omega_h$,
   (iii) $\underline{\boldsymbol{P}}_A \underline{\boldsymbol{\Pi}} = 0$,

where $\boldsymbol{P}$ and $\underline{\boldsymbol{P}}_A$ are the L^2 orthogonal projections into $\boldsymbol{W}^h$ and $\underline{\boldsymbol{A}}^h$, respectively. Besides simplifying the analysis considerably, these properties of the projection tells us how to construct an appropriately small $\delta\underline{\boldsymbol{V}}(K)$, so the question of finding such a $\underline{\boldsymbol{\Pi}}$ and the design of the new finite element are linked.

To answer this question, we begin by noting that the Raviart-Thomas projection applied row-wise, which we denote by $\underline{\boldsymbol{\Pi}}_{\mathrm{RT}}$, will satisfy the first two conditions if $\underline{\boldsymbol{V}}(K)$ is set to $\underline{\boldsymbol{RT}}^k(K)$. Recall that the projection $\underline{\boldsymbol{\Pi}}_{\mathrm{RT}}$ is determined by the equations

$$\langle(\underline{\boldsymbol{\Pi}}_{\mathrm{RT}}\boldsymbol{\sigma})\boldsymbol{n},\boldsymbol{\mu}\rangle_F = \langle\boldsymbol{\sigma}\boldsymbol{n},\boldsymbol{\mu}\rangle_F, \quad \text{for all } \boldsymbol{\mu}\in\boldsymbol{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \quad (1.6a)$$

$$(\underline{\boldsymbol{\Pi}}_{\mathrm{RT}}\boldsymbol{\sigma},\underline{\boldsymbol{v}})_K = (\boldsymbol{\sigma},\underline{\boldsymbol{v}})_K, \qquad \text{for all } \underline{\boldsymbol{v}}\in\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K), \quad (1.6b)$$

where integration over K is indicated by $(\cdot,\cdot)_K$ and integration over F by $\langle\cdot,\cdot\rangle_F$. Hence, any projection $\underline{\boldsymbol{\Pi}}$ *determined* by the equations

$$\langle(\underline{\boldsymbol{\Pi}}\boldsymbol{\sigma})\boldsymbol{n},\boldsymbol{\mu}\rangle_F = \langle\boldsymbol{\sigma}\boldsymbol{n},\boldsymbol{\mu}\rangle_F, \qquad \text{for all } \boldsymbol{\mu}\in\boldsymbol{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \quad (1.7a)$$

$$(\underline{\boldsymbol{\Pi}}\boldsymbol{\sigma},\underline{\boldsymbol{v}})_K = (\boldsymbol{\sigma},\underline{\boldsymbol{v}})_K, \qquad \text{for all } \underline{\boldsymbol{v}}\in\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K), \quad (1.7b)$$

$$(\underline{\boldsymbol{\Pi}}\boldsymbol{\sigma},\underline{\boldsymbol{\eta}})_K = (\boldsymbol{\sigma},\underline{\boldsymbol{\eta}})_K, \qquad \text{for all } \underline{\boldsymbol{\eta}}\in\underline{\tilde{\boldsymbol{A}}}^k(K), \quad (1.7c)$$

where

$$\underline{\tilde{\boldsymbol{A}}}^k(K) := \underline{\boldsymbol{A}}^k(K)\cap\tilde{\boldsymbol{\mathcal{P}}}^k(K), \quad (1.7d)$$

would immediately satisfy the three above-mentioned conditions (i)–(iii). Regarding the notations above, in (1.7d), $\underline{\boldsymbol{A}}^k(K)$ is as in (1.5b). Throughout, $\boldsymbol{n}$ is used to generically denote the unit outward normal (always a column vector) of a domain. The domain will be clear from the context, e.g., in (1.3a) it is Ω, while in (1.7a), it is the outward normal of K.

The main contribution of this paper is to show that such a projection as in (1.7) can be constructed for $k\geq 1$ if we set

$$\delta\underline{\boldsymbol{V}}(K) := \operatorname{curl}(\operatorname{curl}(\underline{\tilde{\boldsymbol{A}}}^k(K))\,\underline{\boldsymbol{b}}_K),$$

where $\underline{\boldsymbol{b}}_K$ is the *symmetric bubble matrix* defined by

$$\underline{\boldsymbol{b}}_K := \sum_{\ell=0}^{3}\lambda_{\ell-3}\lambda_{\ell-2}\lambda_{\ell-1}\,(\operatorname{grad}\lambda_\ell)^t\operatorname{grad}\lambda_\ell. \quad (1.8)$$

Here and throughout, the indices on barycentric coordinates are always calculated mod 4. We note that $(\operatorname{grad}\lambda_\ell)^t\operatorname{grad}\lambda_\ell$ is a 3×3 matrix since, following our convention, $\operatorname{grad}\lambda_\ell$ is a row vector. Let us emphasize that whereas the scalar bubble b_K is a polynomial of degree four, each of the entries of $\underline{\boldsymbol{b}}_K$ is a polynomial of degree three. Moreover, whereas the scalar bubble vanishes on ∂K, the matrix bubble $\underline{\boldsymbol{b}}_K$ *does not*: only the tangential components of its rows vanish. Note that our family of elements does not include the case $k=0$ because in that case $\delta\underline{\boldsymbol{V}}(K)$ is the empty set.

Let us briefly describe the idea behind our choice of the space $\delta\underline{\boldsymbol{V}}(K)$. Comparing the equations defining the projections $\underline{\boldsymbol{\Pi}}^{\mathrm{RT}}$ and $\underline{\boldsymbol{\Pi}}$, we see that to obtain the smallest $\underline{\boldsymbol{V}}(K)$, we should add *just enough* functions to $\underline{\boldsymbol{RT}}^k(K)$ to enable enforcement of weak symmetry, without changing the ranges of the normal trace and the divergence (note that the first two equations defining the projections determine the divergence). Thus we ask: What are the matrix-functions $\underline{\boldsymbol{\psi}}$ in $\underline{\boldsymbol{\mathcal{P}}}^{k+1}(K)$ which satisfy $\operatorname{div}\underline{\boldsymbol{\psi}}|_K\equiv 0$ and $\underline{\boldsymbol{\psi}}\,\boldsymbol{n}|_{\partial K}\equiv 0$? We show later (in Lemma 2.4) that $\underline{\boldsymbol{\psi}}$ must be of the form $\underline{\boldsymbol{\psi}} = \operatorname{curl}(\operatorname{curl}(\underline{\boldsymbol{v}})\,\underline{\boldsymbol{b}}_K)$ for some $\underline{\boldsymbol{v}}\in\underline{\boldsymbol{\mathcal{P}}}^k(K)$. Here the bubble matrix of (1.8) appears.

The definition of the bubble matrix is inspired by [13]. To motivate, we make two observations regarding vector functions $\boldsymbol{v}$ in $\boldsymbol{\mathcal{P}}^{k+1}(K)$ satisfying $\operatorname{div} \boldsymbol{v}|_K \equiv 0$ and $\boldsymbol{v} \cdot \boldsymbol{n}|_{\partial K} \equiv 0$. The first observation is that the results in [13] show that $\boldsymbol{v} = \operatorname{curl}(p_{k-1}\boldsymbol{w}_{ij})$ for some $p_{k-1} \in \mathcal{P}^{k-1}(K)$, where $\boldsymbol{w}_{ij}$ is a generalized Whitney-type basis function of the form

$$\boldsymbol{w}_{ij} = \lambda_{i-3}\lambda_{i-2}\lambda_{i-1}\operatorname{grad}\lambda_i - \lambda_{j-3}\lambda_{j-2}\lambda_{j-1}\operatorname{grad}\lambda_j.$$

The second is the obvious observation that matrix multiplication of the sum in (1.8) with the vector crossproduct $\operatorname{grad}\lambda_k \times \operatorname{grad}\lambda_l$ (where k and l are indices different from i and j) generates a vector collinear to $\boldsymbol{w}_{ij}$. Combining these observations, we find that the above mentioned vector $\boldsymbol{v}$ can be expressed as $\operatorname{curl}(\boldsymbol{r}\underline{\boldsymbol{b}}_K)$ for some $\boldsymbol{r}$ in $\boldsymbol{\mathcal{P}}^{k-1}(K)$, and thus the bubble matrix $\underline{\boldsymbol{b}}_K$ naturally presents itself.

1.3. **Analysis of the new method.** To briefly compare our analysis to Stenberg's [18], we begin by noting that our space $\underline{\boldsymbol{V}}(K)$ is strictly contained in the space that he used for his Raviart-Thomas version. Our choice of $\underline{\boldsymbol{V}}(K)$ allows us to use the projection $\underline{\boldsymbol{\Pi}}$ to carry out a simple error analysis. In particular, we have no need for the macro-element technique used in [18]. Of possibly lesser importance are a few other improvements we make: Optimal estimates for the approximation to $\underline{\boldsymbol{\rho}}$ are proved in [18] using quasi-uniform meshes, whereas here we are able to prove estimates without such an assumption. Note that in order to obtain optimal estimates for $\boldsymbol{\rho}$, we have made critical use of the recent breakthrough of [5]; see also [12, 8]. Finally, we note that our analysis yields a superconvergence result for a new displacement approximation on the element boundaries arising from hybridization.

1.4. **Organization of the paper.** In the next section, we give the new finite element for weakly symmetric stresses. In Section 3, we investigate the mixed method obtained using the new element and show that, by hybridizing it, the only globally coupled degrees of freedom of the method are those of a displacement approximation defined on the element boundaries. In Section 4, we obtain an estimate of the error in the stress by using a typical energy argument. In Section 5, we use a duality argument to estimate the error in the displacement. In Section 6, we prove optimal estimates for the rotation $\underline{\boldsymbol{\rho}}$. In Section 7 , we obtain estimates for the approximation to the displacement on the element boundaries. We end in Section 8, briefly examining a postprocessing scheme that yields a better approximation of the displacement.

2. A NEW FINITE ELEMENT FOR STRESS TENSORS

2.1. **Definition and unisolvency.** Adopting the Ciarlet [10] style of finite element definition, we denote the new finite element by the triple $(K, \underline{\boldsymbol{V}}(K), \Sigma(K))$, whose components are defined as follows. The geometric object K is a simplex. The set of degrees of freedom $\Sigma(K)$ are the following linear functionals

$$\ell_{\boldsymbol{\mu}}(\underline{\boldsymbol{\sigma}}) = \langle \underline{\boldsymbol{\sigma}}\boldsymbol{n}, \boldsymbol{\mu}\rangle_F, \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F), \text{ for all faces } F \text{ of } K, \tag{2.9a}$$

$$\ell_{\underline{\boldsymbol{v}}}(\underline{\boldsymbol{\sigma}}) = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_K, \quad \text{for all } \underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K), \tag{2.9b}$$

$$\ell_{\underline{\boldsymbol{\eta}}}(\underline{\boldsymbol{\sigma}}) = (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\eta}})_K, \quad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\tilde{\boldsymbol{A}}}^k(K), \tag{2.9c}$$

where $\underline{\tilde{A}}^k(K)$ is given by (1.7d). Typically, one would use specific μ, $\underline{v}$, and η that form a basis for $\mathcal{P}^k(F)$, $\underline{\mathcal{P}}^{k-1}(K)$, and $\underline{\tilde{A}}^k(K)$, respectively, to obtain a set of linearly independent functionals.

The space $\underline{V}(K)$ is given by

$$\underline{V}(K) := \underline{V}^k(K) := \underline{RT}^k(K) + B(\underline{\tilde{A}}^k(K)), \qquad (2.10a)$$

where

$$B(\underline{\eta}) := \begin{cases} \operatorname{curl}\left(\operatorname{curl}\left(\underline{\eta}\right) b_K\right) & \text{if } d = 2, \\ \operatorname{curl}\left(\operatorname{curl}\left(\underline{\eta}\right) \underline{b}_K\right) & \text{if } d = 3. \end{cases} \qquad (2.10b)$$

Here b_K is the scalar bubble of the triangle K (equaling the product of all barycentric coordinates) and $\underline{b}_K$ is the symmetric bubble matrix defined by (1.8). In the three-dimensional case, curl acts row-wise as per our convention, but in the two-dimensional case, curl of matrices and column vectors are defined by

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

respectively. The following theorem is the main result of this section.

Theorem 2.1 (Unisolvency). *Let K be a triangle or a tetrahedron. Any $\underline{\sigma}$ in $\underline{V}^k(K)$ is uniquely determined by the above degrees of freedom.*

The remainder of this section is devoted to a proof of this theorem. We begin with a simple lemma on the following form generated by the bubbles:

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

Lemma 2.2. *The form $(\!(\underline{u}, \underline{v})\!)_b$ is an inner product on $\underline{\mathcal{P}}^k(K)$.*

Proof. The result is obvious when $d = 2$. For $d = 3$, recalling the definition of $\underline{b}_K$, and simplifying the Frobenius product,

$$(\underline{v}\, \underline{b}_K, \underline{v})_K = \sum_{i=0}^{3} (\lambda_{i-3}\lambda_{i-2}\lambda_{i-1}\, \underline{v}(\operatorname{grad}\lambda_i)^t, \underline{v}(\operatorname{grad}\lambda_i)^t)_K,$$

Since $\lambda_{i-3}\lambda_{i-2}\lambda_{i-1} \geq 0$ on K, all the terms in the sum above are non-negative. Hence, whenever $(\!(\underline{v}, \underline{v})\!)_b = 0$, the matrix $\underline{v}$ must have $(\operatorname{grad}\lambda_i)^t$ (for all $i = 0, \ldots, 3$) in its kernel. Thus, $\underline{v}$ vanishes on K. The symmetry of $(\!(\underline{u}, \underline{v})\!)_b$ is easy to see, so the proof is complete. $\square$

Lemma 2.3. *For any matrix $\underline{u}$, all the row vectors of the products $\underline{u}\, \underline{b}_K$ and $\underline{u}\, b_K$ have vanishing tangential trace on ∂K.*

Proof. Since the statement is obvious for $\underline{u}\, b_K$, we only consider the matrix product

$$\underline{u}\, \underline{b}_K = \sum_{i=0}^{3} \lambda_{i-3}\lambda_{i-2}\lambda_{i-1}\, \underline{u}(\operatorname{grad}\lambda_i)^t \operatorname{grad}\lambda_i.$$

Let F_l denote the face of K where λ_l vanishes (for any index l in $\{0, 1, 2, 3\}$). The trace of $\underline{u}\, \underline{b}_K$ on F_l is

$$\underline{u}\, \underline{b}_K|_{F_l} = \lambda_{l-3}\lambda_{l-2}\lambda_{l-1}\, \underline{u}\, (\operatorname{grad}\lambda_l)^t\, \operatorname{grad}\lambda_l|_{F_l}.$$

Therefore, all the row vectors of $\underline{\boldsymbol{u}}\,\underline{\boldsymbol{b}}_K|_{F_l}$ are collinear to $\operatorname{grad} \lambda_l$, and consequently has vanishing tangential components on F_l. Since this holds for any face F_l, we have proved the lemma. $\qquad\square$

Lemma 2.4. *Any $\underline{\boldsymbol{\psi}} \in \underline{\boldsymbol{RT}}^k(K)$ satisfies $\operatorname{div}\underline{\boldsymbol{\psi}} = 0$ and $\underline{\boldsymbol{\psi}}\,\boldsymbol{n}|_{\partial K} = 0$ if and only if there exists a $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ such that $\underline{\boldsymbol{\psi}} = B(\underline{\boldsymbol{v}})$.*

Proof. If $\underline{\boldsymbol{\psi}} = B(\underline{\boldsymbol{v}})$ for some $\underline{\boldsymbol{v}}$ in $\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ then it is obvious that $\underline{\boldsymbol{\psi}}$ is in $\underline{\boldsymbol{RT}}^k(K)$ and $\operatorname{div}\underline{\boldsymbol{\psi}} = 0$. Moreover, when $d = 2$, the components of $B(\underline{\boldsymbol{v}})\boldsymbol{n}|_{\partial K}$ equal the tangential derivatives along ∂K of the components of $(\operatorname{curl}\underline{\boldsymbol{v}})b_K$, which vanish on ∂K by Lemma 2.3.

That $B(\underline{\boldsymbol{v}})\boldsymbol{n}|_{\partial K} = 0$ is also true when $d = 3$. Indeed, if $\boldsymbol{v}_i$ is the ith row vector of $\underline{\boldsymbol{v}}$, then by a vector identity, the ith component of $B(\underline{\boldsymbol{v}})\boldsymbol{n}|_{\partial K}$ equals

$$\boldsymbol{n} \cdot \operatorname{curl}\left(\operatorname{curl}(\boldsymbol{v}_i)\underline{\boldsymbol{b}}_K\right) = \operatorname{div}_\tau(\operatorname{curl}(\boldsymbol{v}_i)\underline{\boldsymbol{b}}_K \times \boldsymbol{n}),$$

where div_τ denotes the surface divergence and "$\times$" denotes the vector cross product in three dimensions. The argument of div_τ above is zero, because by Lemma 2.3, all the tangential components of $\operatorname{curl}(\boldsymbol{v}_i)\underline{\boldsymbol{b}}_K$ are zero. Thus we have proved that $\underline{\boldsymbol{\psi}} = B(\underline{\boldsymbol{v}})$ satisfies all the stated properties (for $d = 2$ and 3).

Let us now prove the converse. Suppose we are given a $\underline{\boldsymbol{\psi}}$ in $\underline{\boldsymbol{RT}}^k(K)$ such that $\operatorname{div}\underline{\boldsymbol{\psi}} = 0$ on K and $\underline{\boldsymbol{\psi}}\boldsymbol{n} = 0$ on ∂K. To construct the required $\underline{\boldsymbol{v}}$, consider the problem of finding a $\underline{\boldsymbol{v}}$ in

$$\underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K) \equiv \{\underline{\boldsymbol{w}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K) : (\operatorname{grad}\boldsymbol{z}, \underline{\boldsymbol{w}})_K = 0 \text{ for all } \boldsymbol{z} \in \boldsymbol{\mathcal{P}}^k(K)\}$$

such that

$$(\!(\operatorname{curl}\underline{\boldsymbol{v}}, \operatorname{curl}\underline{\boldsymbol{\omega}})\!)_b = (\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\omega}})_K, \tag{2.11}$$

for all $\underline{\boldsymbol{\omega}} \in \underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K)$. That this equation is uniquely solvable for $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K)$ is a consequence of a well known exact sequence property [14], namely the range of $\operatorname{grad} : \underline{\boldsymbol{\mathcal{P}}}^k(K) \mapsto \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ equals the kernel of curl in $\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$. If the right hand side of (2.11) vanishes, then by Lemma 2.2, $\operatorname{curl}\underline{\boldsymbol{v}} = 0$, so $\underline{\boldsymbol{v}}$ is in the kernel of curl. But at the same time, $\underline{\boldsymbol{v}}$ being in $\underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K)$, is in the orthogonal complement of the range of grad, hence $\underline{\boldsymbol{v}}$ vanishes. This shows that (2.11) has a unique solution.

Having established that there is a unique $\underline{\boldsymbol{v}}$ in $\underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K) \subseteq \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ satisfying (2.11), the proof will be completed if we show that the solution $\underline{\boldsymbol{v}}$ satisfies

$$\underline{\boldsymbol{\psi}} = B(\underline{\boldsymbol{v}}). \tag{2.12}$$

To this end, first observe that by the definition of $\underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K)$, we can split any $\underline{\boldsymbol{\xi}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ as $\underline{\boldsymbol{\xi}} = \operatorname{grad}\boldsymbol{q} + \underline{\boldsymbol{\omega}}$, with $\boldsymbol{q}$ in $\boldsymbol{\mathcal{P}}^k(K)$ and $\underline{\boldsymbol{\omega}}$ in $\underline{\boldsymbol{\mathcal{Z}}}^{k-1}(K)$. Then

$$(\!(\operatorname{curl}\underline{\boldsymbol{v}}, \operatorname{curl}\underline{\boldsymbol{\xi}})\!)_b = (\!(\operatorname{curl}\underline{\boldsymbol{v}}, \operatorname{curl}\underline{\boldsymbol{\omega}})\!)_b = (\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\omega}})_K = (\underline{\boldsymbol{\psi}}, \operatorname{grad}\boldsymbol{q} + \underline{\boldsymbol{\omega}})_K = (\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\xi}})_K,$$

where we have used the fact that $\operatorname{div}\underline{\boldsymbol{\psi}} = 0$ and $\underline{\boldsymbol{\psi}}\boldsymbol{n}|_{\partial K} = 0$ imply $(\underline{\boldsymbol{\psi}}, \operatorname{grad}\boldsymbol{q})_K = 0$. Thus (2.11) in fact holds for all $\underline{\boldsymbol{\omega}}$ in $\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$. Integrating the left hand side of (2.11) by parts using Lemma 2.3, we get

$$(B(\underline{\boldsymbol{v}}), \underline{\boldsymbol{\omega}})_K = (\underline{\boldsymbol{\psi}}, \underline{\boldsymbol{\omega}}) \quad \text{for all } \underline{\boldsymbol{\omega}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K).$$

Letting $\boldsymbol{J}_i$ denote the i-th row of $B(\underline{\boldsymbol{v}}) - \underline{\boldsymbol{\psi}}$, we therefore have

$$(\boldsymbol{J}_i, \boldsymbol{\omega})_K = 0, \quad \text{for all } \boldsymbol{\omega} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{2.13a}$$

$$\langle \boldsymbol{J}_i \cdot \boldsymbol{n}, \mu \rangle_F = 0, \quad \text{for all } \mu \in \mathcal{P}^k(F), \tag{2.13b}$$

and for all faces F of K. Note that to obtain (2.13b), we used the consequence of Lemma 2.3 that $B(\underline{\boldsymbol{v}})\boldsymbol{n}|_{\partial K} = 0$ and the fact that $\underline{\boldsymbol{\psi}}\boldsymbol{n}|_{\partial K} = 0$. Since $\boldsymbol{J}_i$ is in $\boldsymbol{RT}^k(K)$, by the well known [16] degrees of freedom of the Raviart-Thomas space, the equations of (2.13) imply that all the degrees of freedom of $\boldsymbol{J}_i$ are zero. Hence $\boldsymbol{J}_i$ (for all i) vanishes and we have proved (2.12).   □

Lemma 2.5. *If $B(\underline{\boldsymbol{\eta}}) = \underline{\boldsymbol{\psi}}$ with $\underline{\boldsymbol{\psi}} \in \underline{\boldsymbol{RT}}^k(K)$ and $\underline{\boldsymbol{\eta}} \in \underline{\tilde{\boldsymbol{A}}}^k(K)$ then $\underline{\boldsymbol{\eta}} = 0$.*

Proof. We note that $\operatorname{div}\underline{\boldsymbol{\psi}} = 0$ on K and $\underline{\boldsymbol{\psi}}\,\boldsymbol{n} = 0$ on ∂K. By Lemma 2.4 we have that $\underline{\boldsymbol{\psi}} = B(\underline{\boldsymbol{v}})$ for some $\underline{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, and hence $B(\underline{\boldsymbol{\eta}}) = B(\underline{\boldsymbol{v}})$. Applying integration by parts and using Lemma 2.3 and Lemma 2.2 we can easily show that $\operatorname{curl}(\underline{\boldsymbol{\eta}}) = \operatorname{curl}(\underline{\boldsymbol{v}})$. We first complete the proof in the three-dimensional case. Since $\underline{\boldsymbol{\eta}} \in \underline{\tilde{\boldsymbol{A}}}^k(K)$ we have

$$\underline{\boldsymbol{\eta}} = a(\boldsymbol{z}) \equiv \begin{pmatrix} 0 & z_3 & -z_2 \\ -z_3 & 0 & z_1 \\ z_2 & -z_1 & 0 \end{pmatrix}$$

for some $\boldsymbol{z} = (z_1, z_2, z_3)^t \in \tilde{\boldsymbol{\mathcal{P}}}^k(K)$. By virtue of the identity

$$\operatorname{curl} a(\boldsymbol{z}) = (\operatorname{grad} \boldsymbol{z})^t - (\operatorname{div} \boldsymbol{z})\underline{\boldsymbol{I}}, \tag{2.14}$$

we have that

$$\operatorname{curl}(\underline{\boldsymbol{v}}) : \underline{\boldsymbol{I}} = ((\operatorname{grad} \boldsymbol{z})^t - (\operatorname{div} \boldsymbol{z})\underline{\boldsymbol{I}}) : \underline{\boldsymbol{I}} = -2 \operatorname{div} \boldsymbol{z},$$

i.e., $\operatorname{div} \boldsymbol{z} \in \boldsymbol{\mathcal{P}}^{k-2}(K)$. Therefore, since $(\operatorname{grad} \boldsymbol{z})^t = \operatorname{curl}(\underline{\boldsymbol{v}}) + (\operatorname{div} \boldsymbol{z})\underline{\boldsymbol{I}}$ we have that $\operatorname{grad} \boldsymbol{z} \in \boldsymbol{\mathcal{P}}^{k-2}(K)$. This implies that $\boldsymbol{z} \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, but since $\boldsymbol{\mathcal{P}}^{k-1}(K) \cap \tilde{\boldsymbol{\mathcal{P}}}^k(K) = \{0\}$ for $k \geq 1$, we have that $\boldsymbol{z} = 0$. This shows that $\underline{\boldsymbol{\eta}} = 0$.

In the two-dimensional case,

$$\underline{\boldsymbol{\eta}} = \begin{pmatrix} 0 & z \\ -z & 0 \end{pmatrix},$$

for some $z \in \tilde{\mathcal{P}}^k(K)$. Using that $\operatorname{curl}(\underline{\boldsymbol{\eta}}) = (\operatorname{grad} z)^t$ we can finish the proof as in the three-dimensional case.   □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. We first claim that $\dim(\underline{\boldsymbol{V}}^k(K))$ equals the number of independent degrees of freedom. To calculate $\dim(\underline{\boldsymbol{V}}^k(K))$, we first note that by Lemma 2.5 the two spaces appearing in (2.10a) are linearly independent. Lemma 2.5 also gives us that $B : \underline{\tilde{\boldsymbol{A}}}^k(K) \longmapsto \underline{\boldsymbol{V}}^k(K)$ is one-one. Hence,

$$\dim(\underline{\boldsymbol{V}}^k(K)) = \dim(\underline{\boldsymbol{RT}}^k(K)) + \dim(\underline{\tilde{\boldsymbol{A}}}^k(K)).$$

It is standard [16, 17] to show that the number of degrees of freedom in (2.9a) and (2.9b) equal $\dim(\underline{\boldsymbol{RT}}^k(K))$. The number of degrees of freedom in (2.9c) obviously equal $\dim(\underline{\tilde{\boldsymbol{A}}}^k(K))$ so the claim is verified.

Next, we prove that if $\underline{\boldsymbol{\sigma}}$ is a function in $\underline{\boldsymbol{V}}^k(K)$ with vanishing degrees of freedom, then $\underline{\boldsymbol{\sigma}}$ vanishes. Standard arguments [16, 17] show that whenever

$$\ell_\mu(\underline{\boldsymbol{\sigma}}) = \ell_{\underline{\boldsymbol{v}}}(\underline{\boldsymbol{\sigma}}) = 0$$

for all $\boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F)$ for all faces F of K and all $\underline{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, we have

$$\operatorname{div} \underline{\boldsymbol{\sigma}} = 0 \quad \text{and} \quad \underline{\boldsymbol{\sigma}}\boldsymbol{n}|_{\partial K} = 0. \tag{2.15}$$

By the definition of the space $\underline{\boldsymbol{V}}^k(K)$, this implies that $\boldsymbol{\sigma} = B(\boldsymbol{\zeta}) + \boldsymbol{\psi}$, for some $\boldsymbol{\zeta} \in \underline{\tilde{\boldsymbol{A}}}^k(K)$ and some $\boldsymbol{\psi}$ in $\underline{\boldsymbol{RT}}^k(K)$ with $\operatorname{div}\boldsymbol{\psi} = 0$ and $\boldsymbol{\psi}\boldsymbol{n}|_{\partial K} = 0$. Using Lemma 2.4, we therefore conclude that there is a $\underline{\boldsymbol{w}}$ in $\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)$ such that

$$\boldsymbol{\sigma} = B(\boldsymbol{\zeta} + \underline{\boldsymbol{w}}).$$

Then, using the last set of degrees of freedom, namely that $\ell_{\boldsymbol{\eta}}(\boldsymbol{\sigma}) = 0$ for $\boldsymbol{\eta} \in \underline{\tilde{\boldsymbol{A}}}^k(K)$, and again $\ell_{\underline{\boldsymbol{v}}}(\boldsymbol{\sigma}) = 0$, we obtain

$$\begin{aligned}
0 = (\boldsymbol{\sigma}, \boldsymbol{\eta} + \underline{\boldsymbol{v}})_K &= (B(\boldsymbol{\zeta} + \underline{\boldsymbol{w}}), \boldsymbol{\eta} + \underline{\boldsymbol{v}})_K \\
&= (\!(\operatorname{curl}(\boldsymbol{\zeta} + \underline{\boldsymbol{w}}), \operatorname{curl}(\boldsymbol{\eta} + \underline{\boldsymbol{v}}))\!)_b.
\end{aligned}$$

where we used integration by parts and Lemma 2.3. Choosing $\boldsymbol{\eta} = \boldsymbol{\zeta}$ and $\underline{\boldsymbol{v}} = \underline{\boldsymbol{w}}$, by Lemma 2.2, we find that $\operatorname{curl}(\boldsymbol{\zeta} + \underline{\boldsymbol{w}}) = 0$. Hence, $\boldsymbol{\sigma} = B(\boldsymbol{\zeta} + \underline{\boldsymbol{w}}) = 0$. $\square$

2.2. **Mappings.** For implementation purposes, as well as for theoretical scaling arguments, we need mappings between the finite elements. To define the appropriate mapping for our finite element, let $\hat{K}$ be a fixed 'reference' simplex. It can be mapped homeomorphically to any other simplex K by a mapping of the form $\mathcal{F}(\hat{\boldsymbol{x}}) = \underline{\boldsymbol{M}}\hat{\boldsymbol{x}} + \boldsymbol{b}$. Matrix-valued functions $\hat{\boldsymbol{\sigma}}$ and (column) vector-valued functions $\hat{\boldsymbol{w}}$ on $\hat{K}$ are mapped to functions on K by

$$\boldsymbol{\sigma}(\boldsymbol{x}) = |\det \underline{\boldsymbol{M}}|^{-1}\underline{\boldsymbol{M}}\,\hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{x}})\,\underline{\boldsymbol{M}}^t, \tag{2.16}$$

$$\boldsymbol{w}(\boldsymbol{x}) = \underline{\boldsymbol{M}}^{-t}\,\hat{\boldsymbol{w}}(\hat{\boldsymbol{x}}). \tag{2.17}$$

Additionally, letting $\check{\boldsymbol{\tau}}(\hat{\boldsymbol{x}}) = \underline{\boldsymbol{M}}^t\boldsymbol{\tau}(\boldsymbol{x})\,\underline{\boldsymbol{M}}$, it is easy to see that

$$(\boldsymbol{\sigma}, \boldsymbol{\tau})_K = (\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{\tau}})_{\hat{K}} \tag{2.18a}$$

$$(\boldsymbol{\sigma}, \operatorname{grad}\boldsymbol{w})_K = (\hat{\boldsymbol{\sigma}}, \operatorname{grad}\hat{\boldsymbol{w}})_{\hat{K}}, \tag{2.18b}$$

$$(\operatorname{div}\boldsymbol{\sigma}, \boldsymbol{v})_K = (\operatorname{div}\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{v}})_{\hat{K}}, \tag{2.18c}$$

$$\langle\boldsymbol{\sigma}\boldsymbol{n}, \boldsymbol{\mu}\rangle_F = \langle\hat{\boldsymbol{\sigma}}\hat{\boldsymbol{n}}, \hat{\boldsymbol{\mu}}\rangle_{\hat{F}}, \tag{2.18d}$$

where F is the image of a face $\hat{F}$ under $\mathcal{F}$ and $\hat{\boldsymbol{n}}$ is the unit outward normal of $\hat{K}$. Note that the space $\underline{\boldsymbol{V}}^k(\hat{K})$ under the mapping (2.16) is carried one-one onto $\underline{\boldsymbol{V}}^k(K)$. Furthermore, the degrees of freedom are also mapped over. For instance, by (2.18d),

$$\ell_{\boldsymbol{\eta}}(\boldsymbol{\sigma}) = (\boldsymbol{\sigma}, \boldsymbol{\eta})_K = (\hat{\boldsymbol{\sigma}}, \check{\boldsymbol{\eta}})_{\hat{K}} = \ell_{\check{\boldsymbol{\eta}}}(\hat{\boldsymbol{\sigma}}),$$

and the mapping $\boldsymbol{\eta} \mapsto \check{\boldsymbol{\eta}}$ carries the space of antisymmetric matrices $\underline{\boldsymbol{A}}^p(K)$ one-one onto $\underline{\boldsymbol{A}}^p(\hat{K})$ for every degree p. The next result is proved using these observations.

Corollary 2.6. *Suppose $\boldsymbol{\sigma}$ in $\underline{\boldsymbol{V}}^k(K)$ satisfies*

$$\langle\boldsymbol{\sigma}\boldsymbol{n}, \boldsymbol{\mu}\rangle_F = \langle\boldsymbol{E}, \boldsymbol{\mu}\rangle_F \qquad \textit{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F), \textit{ for all faces } F \textit{ of } K, \tag{2.19a}$$

$$(\boldsymbol{\sigma}, \underline{\boldsymbol{v}})_K = (\underline{\boldsymbol{G}}, \underline{\boldsymbol{v}})_K \qquad \textit{for all } \underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K), \tag{2.19b}$$

$$(\boldsymbol{\sigma}, \underline{\boldsymbol{\eta}})_K = (\underline{\boldsymbol{H}}, \underline{\boldsymbol{\eta}})_K \qquad \textit{for all } \underline{\boldsymbol{\eta}} \in \underline{\tilde{\boldsymbol{A}}}^k(K). \tag{2.19c}$$

Then there is a constant $C > 0$ depending only on k and the shape regularity of K such that

$$C\|\boldsymbol{\sigma}\|_{L^2(K)} \leq h_K^{1/2}\|\boldsymbol{E}\|_{L^2(\partial K)} + \|\underline{\boldsymbol{G}}\|_{L^2(K)} + \|\underline{\boldsymbol{H}}\|_{L^2(K)}$$

where $h_K = \operatorname{diam}(K)$ and $\|\cdot\|_X$ denotes the X-norm.

Proof. Map $\underline{\boldsymbol{\sigma}}$ to a fixed reference element $\hat{K}$. On the finite dimensional space $\underline{\boldsymbol{V}}^k(\hat{K})$, because of Theorem 2.1, the norms $\|\hat{\underline{\boldsymbol{\sigma}}}\|_{L^2(\hat{K})}$ and

$$\sup_{\substack{\hat{\boldsymbol{\mu}}\in\boldsymbol{\mathcal{P}}^k(\hat{F})\\ \hat{F}\subseteq\partial\hat{K}}} \frac{\langle\hat{\underline{\boldsymbol{\sigma}}}\hat{\boldsymbol{n}},\hat{\boldsymbol{\mu}}\rangle_{\hat{F}}}{\|\hat{\boldsymbol{\mu}}\|_{L^2(\hat{F})}} + \sup_{\check{\underline{\boldsymbol{\tau}}}\in\underline{\boldsymbol{\mathcal{P}}}^{k-1}(\hat{K})} \frac{(\hat{\underline{\boldsymbol{\sigma}}},\check{\underline{\boldsymbol{\tau}}})_{\hat{K}}}{\|\check{\underline{\boldsymbol{\tau}}}\|_{L^2(\hat{K})}} + \sup_{\check{\underline{\boldsymbol{\eta}}}\in\check{\underline{\boldsymbol{A}}}^k(\hat{K})} \frac{(\hat{\underline{\boldsymbol{\sigma}}},\check{\underline{\boldsymbol{\eta}}})_{\hat{K}}}{\|\check{\underline{\boldsymbol{\eta}}}\|_{L^2(\hat{K})}} \qquad (2.20)$$

are equivalent norms. Let us write $A \sim B$ whenever there are positive constants C_1 and C_2 depending only on the shape regularity of K such that $C_1 A \leq B \leq C_2 A$. If $\boldsymbol{\mu}$ is mapped as in (2.17) and $\underline{\boldsymbol{\sigma}}$ as in (2.16), then it is easy to establish the following two sided estimates (by standard arguments, see e.g. [10]):

$$\|\underline{\boldsymbol{\sigma}}\|^2_{L^2(K)} \sim h_K^{4-d}\|\hat{\underline{\boldsymbol{\sigma}}}\|^2_{L^2(\hat{K})}$$

$$\|\underline{\boldsymbol{\tau}}\|^2_{L^2(K)} \sim h_K^{d-4}\|\check{\underline{\boldsymbol{\tau}}}\|^2_{L^2(\hat{K})}$$

$$\|\boldsymbol{\mu}\|^2_{L^2(F)} \sim h_K^{d-3}\|\hat{\boldsymbol{\mu}}\|^2_{L^2(\hat{F})}.$$

Applying these for the denominators in (2.20), and applying the identities (2.18a) and (2.18d) for the numerators, we find that

$$C\|\underline{\boldsymbol{\sigma}}\|_{L^2(K)} \leq \sup_{\substack{\boldsymbol{\mu}\in\boldsymbol{\mathcal{P}}^k(F)\\ F\subseteq\partial K}} \frac{h_K^{1/2}\langle\boldsymbol{E},\boldsymbol{\mu}\rangle_F}{\|\boldsymbol{\mu}\|_{L^2(F)}} + \sup_{\underline{\boldsymbol{\tau}}\in\underline{\boldsymbol{\mathcal{P}}}^{k-1}(K)} \frac{(\underline{\boldsymbol{G}},\underline{\boldsymbol{\tau}})_K}{\|\underline{\boldsymbol{\tau}}\|_{L^2(K)}} + \sup_{\underline{\boldsymbol{\eta}}\in\check{\underline{\boldsymbol{A}}}^k(K)} \frac{(\underline{\boldsymbol{H}},\underline{\boldsymbol{\eta}})_K}{\|\underline{\boldsymbol{\eta}}\|_{L^2(K)}}$$

and the corollary follows. $\qquad\square$

2.3. **The projection.** Any finite element has a projection (also known as an interpolation operator) [10]. For a smooth $\underline{\boldsymbol{\sigma}}$, its projection into the finite element space $\underline{\boldsymbol{V}}^k(K)$, denoted by $\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}}$, is defined as usual, by

$$\ell(\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}}) = 0$$

for all the degrees of freedom $\ell \in \{\ell_{\boldsymbol{\mu}},\ell_{\underline{\boldsymbol{v}}},\ell_{\underline{\boldsymbol{\eta}}}\}$ defined in (2.9). Note that these equations are the same as (1.7). In other words, the interpolation operator of our finite element is precisely the projection $\underline{\boldsymbol{\Pi}}$ that we sought in (1.7).

Note that $\underline{\boldsymbol{\Pi}}$ can only be applied to $\underline{\boldsymbol{\sigma}}$ that is smooth enough for the right hand sides of (1.7) to make sense. While (1.7b) and (1.7c) make sense for any $\underline{\boldsymbol{\sigma}}$ in $\underline{\boldsymbol{L}}^2(K)$, for (1.7a), we need $\underline{\boldsymbol{\sigma}}$ to have slightly better regularity properties, such as, for example, $\underline{\boldsymbol{\sigma}}$ in $\underline{\boldsymbol{H}}(\operatorname{div},K)\cap\underline{\boldsymbol{L}}^p(K)$ for any $p > 2$ (by the same arguments as for the Raviart-Thomas projection [9]). So as not to get into technicalities concerning the maximal domain of $\underline{\boldsymbol{\Pi}}$, we will tacitly assume that all functions to which we apply $\underline{\boldsymbol{\Pi}}$ are in its domain, which we simply denote by dom $\underline{\boldsymbol{\Pi}}$, a set that satisfies the inclusions

$$\underline{\boldsymbol{H}}(\operatorname{div},K)\cap\underline{\boldsymbol{L}}^p(K) \;\subset\; \operatorname{dom}\underline{\boldsymbol{\Pi}} \;\subset\; \underline{\boldsymbol{H}}(\operatorname{div},K)$$

for $p > 2$. That for all $\underline{\boldsymbol{\sigma}}$ in dom $\underline{\boldsymbol{\Pi}}$, the left hand sides of (1.7) uniquely define $\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}}$ in $\underline{\boldsymbol{V}}^k(K)$, follows from Theorem 2.1. The important properties of $\underline{\boldsymbol{\Pi}}$ that we shall use later are proved in the next theorem. We use C to denote a generic constant, whose values at different occurrences may differ, but will always remain independent of the size of mesh elements.

Theorem 2.7. *Suppose $\underline{\boldsymbol{\sigma}}$ is in* dom $\underline{\boldsymbol{\Pi}}$. *Then*

$$\operatorname{div}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}}) = \boldsymbol{P}(\operatorname{div}\underline{\boldsymbol{\sigma}}), \qquad (2.21)$$

where $\boldsymbol{P}$ is the $L^2(\Omega)$-orthogonal projection onto $\boldsymbol{W}^h$. If in addition, $\underline{\boldsymbol{\sigma}}$ is in $\underline{\boldsymbol{H}}^r(K)$ for some r satisfying $1 \leq r \leq k+1$, then

$$\|\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(K)} \leq C\,h_K^r |\underline{\boldsymbol{\sigma}}|_{H^r(K)}. \tag{2.22}$$

Proof. Let us prove the identity (2.21). For any $\boldsymbol{v}$ in $\boldsymbol{W}^h$, we have

$$\begin{aligned}
(\operatorname{div}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}}), \boldsymbol{v})_K &= -(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}}, \operatorname{grad}\boldsymbol{v})_K + \langle \underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}}\,\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K} \\
&= -(\underline{\boldsymbol{\sigma}}, \operatorname{grad}\boldsymbol{v})_K + \langle \underline{\boldsymbol{\sigma}}\,\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K} \qquad \text{by (1.7b) and (1.7a)}, \\
&= (\operatorname{div}\underline{\boldsymbol{\sigma}}, \boldsymbol{v})_K,
\end{aligned}$$

which readily implies the identity (2.21).

To prove the approximation estimate, we use the triangle inequality

$$\|\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(K)} \leq \|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}}_{\text{RT}}\underline{\boldsymbol{\sigma}}\|_{L^2(K)} + \|\underline{\boldsymbol{\delta}}\|_{L^2(K)}, \tag{2.23}$$

where $\underline{\boldsymbol{\delta}} := (\underline{\boldsymbol{\Pi}} - \underline{\boldsymbol{\Pi}}_{\text{RT}})\underline{\boldsymbol{\sigma}}$ and $\underline{\boldsymbol{\Pi}}_{\text{RT}}$ is the Raviart-Thomas projection defined in (1.6). Note that $\underline{\boldsymbol{\delta}}$ is in $\underline{\boldsymbol{V}}^k(K)$ and satisfies (2.19) with $\boldsymbol{E} = 0, \underline{\boldsymbol{G}} = 0$ and $\underline{\boldsymbol{H}} = \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}}_{\text{RT}}\underline{\boldsymbol{\sigma}}$. Hence by Corollary 2.6,

$$\|\underline{\boldsymbol{\delta}}\|_{L^2(K)} \leq C\,\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}}_{\text{RT}}\underline{\boldsymbol{\sigma}}\|_{L^2(K)}.$$

Using this in (2.23), the estimate of the theorem follows from the well-known approximation properties of the projection $\underline{\boldsymbol{\Pi}}_{\text{RT}}$. $\qquad\square$

2.4. **Inf-sup condition.** Our global finite element spaces are $\underline{\boldsymbol{V}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h$ as defined by (1.3) where the local spaces are given by (1.5) and (2.10). A fundamental property of this triplet of spaces is as follows:

Lemma 2.8. *Given any $\boldsymbol{w}$ in $\boldsymbol{W}^h$ and $\underline{\boldsymbol{\zeta}}$ in $\underline{\boldsymbol{A}}^h$, there exists a $\underline{\boldsymbol{\tau}}$ in $\underline{\boldsymbol{V}}^h$ satisfying*

$$\operatorname{div}\underline{\boldsymbol{\tau}} = \boldsymbol{w},$$
$$(\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\eta}}) = (\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\eta}}), \qquad \forall \underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}^h, \ and$$
$$\|\underline{\boldsymbol{\tau}}\|_{L^2(\Omega)} \leq C\big(\|\boldsymbol{w}\|_{L^2(\Omega)} + \|\underline{\boldsymbol{\zeta}}\|_{L^2(\Omega)}\big),$$

where C only depends on the shape regularity of Ω_h.

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

$$\operatorname{div}\underline{\boldsymbol{\tau}}_{k+1} = \boldsymbol{w},$$
$$(\underline{\boldsymbol{\tau}}_{k+1}, \underline{\boldsymbol{\eta}}) = (\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\eta}}), \qquad \forall \underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}^h, \ and$$
$$\|\underline{\boldsymbol{\tau}}_{k+1}\|_{L^2(\Omega)} \leq C\big(\|\boldsymbol{w}\|_{L^2(\Omega)} + \|\underline{\boldsymbol{\zeta}}\|_{L^2(\Omega)}\big),$$

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 $\underline{\boldsymbol{\Pi}}$, so we may set $\underline{\boldsymbol{\tau}} = \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\tau}}_{k+1}$. Then

$$\begin{aligned}
\operatorname{div}\underline{\boldsymbol{\tau}} &= \boldsymbol{P}\operatorname{div}\underline{\boldsymbol{\tau}}_{k+1} = \boldsymbol{w}, &\text{by Theorem 2.7,} \\
(\underline{\boldsymbol{\tau}}, \underline{\boldsymbol{\eta}}) &= (\underline{\boldsymbol{\tau}}_{k+1}, \underline{\boldsymbol{\eta}}) = (\underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\eta}}), \ \ \forall \underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}^h, &\text{by (1.7b) and (1.7c),}
\end{aligned}$$

which are the first two identities of the lemma. The final estimate of the lemma follows by scaling arguments (cf. Corollary 2.6) and norm equivalences. $\qquad\square$

An immediate consequence of Lemma 2.8 is that the "inf-sup" [9] condition

$$\sup_{\underline{\boldsymbol{\tau}} \in \underline{\boldsymbol{V}}^h} \frac{(\boldsymbol{w}^h, \operatorname{div} \underline{\boldsymbol{\tau}}) + (\boldsymbol{\zeta}^h, \underline{\boldsymbol{\tau}})}{\|\underline{\boldsymbol{\tau}}\|_{H(\operatorname{div}, \Omega)}} \geq C \big(\|\boldsymbol{w}^h\|_{L^2(\Omega)} + \|\boldsymbol{\zeta}^h\|_{L^2(\Omega)} \big) \qquad (2.24)$$

holds for all $\boldsymbol{\zeta}^h$ in $\underline{\boldsymbol{A}}^h$ and $\boldsymbol{w}^h$ in $\boldsymbol{W}^h$. This can be used together with the abstract theory of mixed methods [9] to obtain error estimates for our method. But in later sections, we do a refined error analysis so as to clearly see the dependence of errors in each variable on others.

3. The method and its hybridization

3.1. **Definition.** The method we consider takes the approximation $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h)$ in the space $\underline{\boldsymbol{V}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h$ defined by (1.3) where the local spaces are given by (1.5) and (2.10). It is determined by the equations (1.4).

Proposition 3.1. *There is a unique* $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h) \in \underline{\boldsymbol{V}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h$ *satisfying* (1.4) *for* $k \geq 1$.

Proof. Since (1.4) is a square system we only need to prove uniqueness. To this end, suppose $\boldsymbol{f} = 0$ in (1.4). Setting all test functions to trial functions and adding the equations of (1.4), we obtain that $(\mathcal{A}\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\sigma}}^h) = 0$ which immediately implies that $\underline{\boldsymbol{\sigma}}^h = 0$. Next, choose $\underline{\boldsymbol{\tau}}$ in (1.4a) to be the function given by Lemma 2.8 with $\boldsymbol{w} = \boldsymbol{u}^h$ and $\boldsymbol{\zeta} = \underline{\boldsymbol{\rho}}^h$. Then we see that $\boldsymbol{u}^h$ and $\underline{\boldsymbol{\rho}}^h$ also vanish. $\square$

3.2. **Hybridized formulation.** Implementation of the method in the form (1.4) will result in an indefinite linear system that couples all the three unknowns. Hence we recommend an alternate implementation technique via hybridization that will result in a symmetric positive definite system for a single new variable. We develop this technique now.

The hybridized formulation removes the $\underline{\boldsymbol{H}}(\operatorname{div}, \Omega)$-continuity constraints from the space $\underline{\boldsymbol{V}}^h$, and places them as an additional equation of the method. Accordingly, we need the space

$$\underline{\widetilde{\boldsymbol{V}}}^h = \{\underline{\boldsymbol{v}} : \underline{\boldsymbol{v}}|_K \in \underline{\boldsymbol{V}}^k(K) \text{ for all mesh elements } K \in \Omega_h\}$$

without $\underline{\boldsymbol{H}}(\operatorname{div}, \Omega)$-conformity, as well as a space of Lagrange multipliers

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

The approximate solution given by the hybridized method is $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h) \in \underline{\widetilde{\boldsymbol{V}}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h \times \boldsymbol{M}^h$, satisfying

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{v}})_{\Omega_h} + (\boldsymbol{u}^h, \operatorname{div} \underline{\boldsymbol{v}})_{\Omega_h} + (\underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{v}})_{\Omega_h} + \langle \boldsymbol{\lambda}^h, \underline{\boldsymbol{v}}\boldsymbol{n} \rangle_{\partial\Omega_h} = 0, \qquad (3.25a)$$

$$(\operatorname{div} \underline{\boldsymbol{\sigma}}^h, \boldsymbol{\omega})_{\Omega_h} = (\boldsymbol{f}, \boldsymbol{\omega})_{\Omega_h}, \qquad (3.25b)$$

$$(\underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\eta}})_{\Omega_h} = 0, \qquad (3.25c)$$

$$\langle \underline{\boldsymbol{\sigma}}^h \boldsymbol{n}, \boldsymbol{\mu} \rangle_{\partial\Omega_h} = 0, \qquad (3.25d)$$

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

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

Note that using identities such as in (2.18), the stiffness matrices corresponding to each term in (3.25) can be computed on a fixed reference element and mapped over.

Proposition 3.2. *There is a unique $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h) \in \widetilde{\underline{\boldsymbol{V}}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h \times \boldsymbol{M}^h$ satisfying* (3.25). *Moreover, the first three components of the solution, namely $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h)$ coincide with that of the mixed method* (1.4).

Proof. First observe that if $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h)$ satisfies (3.25), then $\underline{\boldsymbol{\sigma}}^h \in \underline{\boldsymbol{V}}^h$. Moreover, since $\widetilde{\underline{\boldsymbol{V}}}^h \supset \underline{\boldsymbol{V}}^h$, choosing test functions $\underline{\boldsymbol{v}}$ in $\underline{\boldsymbol{V}}^h$, we see that the equations (3.25a)–(3.25c) are identical to the equations of the mixed method (1.4). The second assertion of the proposition is therefore immediate.

To prove that there is a unique $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h)$, since the system (3.25) is square, it is enough to show that the only solution when $\boldsymbol{f} = \boldsymbol{0}$ is the trivial one. When $\boldsymbol{f}$ vanishes, the argument of the previous paragraph and Proposition 3.1 shows that $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h)$ is zero. The equation (3.25a) then becomes

$$\langle \boldsymbol{\lambda}^h, \underline{\boldsymbol{v}}\boldsymbol{n} \rangle_{\partial \Omega_h} = 0 \qquad \text{for all } \underline{\boldsymbol{v}} \in \widetilde{\underline{\boldsymbol{V}}}^h,$$

which readily implies that $\boldsymbol{\lambda}^h = \boldsymbol{0}$. This completes the proof. $\qquad \square$

It is possible to eliminate the variables $\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h$, and $\underline{\boldsymbol{\rho}}^h$ and obtain one global system for $\boldsymbol{\lambda}^h$. Furthermore, once $\boldsymbol{\lambda}^h$ is computed, $\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h$, and $\underline{\boldsymbol{\rho}}^h$ can be computed using $\boldsymbol{\lambda}^h$ locally, element by element. To make such assertions precise, consider the solution of the following problem on *one* element:

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}_{F,G}, \underline{\boldsymbol{v}})_K + (\boldsymbol{u}_{F,G}, \operatorname{div} \underline{\boldsymbol{v}})_K + (\underline{\boldsymbol{\rho}}_{F,G}, \underline{\boldsymbol{v}})_K = G(\underline{\boldsymbol{v}}) \qquad (3.26\mathrm{a})$$

$$(\operatorname{div} \underline{\boldsymbol{\sigma}}_{F,G}, \boldsymbol{\omega})_K = F(\boldsymbol{\omega}) \qquad (3.26\mathrm{b})$$

$$(\underline{\boldsymbol{\sigma}}_{F,G}, \underline{\boldsymbol{\eta}})_K = 0. \qquad (3.26\mathrm{c})$$

Furthermore, let $(\underline{\boldsymbol{\sigma}}_{\boldsymbol{\mu}}, \boldsymbol{u}_{\boldsymbol{\mu}}, \underline{\boldsymbol{\rho}}_{\boldsymbol{\mu}})$ denote the local solution $(\underline{\boldsymbol{\sigma}}_{F,G}, \boldsymbol{u}_{F,G}, \underline{\boldsymbol{\rho}}_{F,G})$ obtained when $G(\underline{\boldsymbol{v}}) = -\langle \boldsymbol{\mu}, \underline{\boldsymbol{v}}\boldsymbol{n} \rangle_{\partial K}$ and $F = 0$, and let $(\underline{\boldsymbol{\sigma}}_{\boldsymbol{f}}, \boldsymbol{u}_{\boldsymbol{f}}, \underline{\boldsymbol{\rho}}_{\boldsymbol{f}})$ denote the local solution $(\underline{\boldsymbol{\sigma}}_{F,G}, \boldsymbol{u}_{F,G}, \underline{\boldsymbol{\rho}}_{F,G})$ obtained when $G = 0$ and $F(\boldsymbol{\omega}) = (\boldsymbol{f}, \boldsymbol{\omega})_K$. Then we have the following theorem.

Theorem 3.3. *The function $(\underline{\boldsymbol{\sigma}}^h, \boldsymbol{u}^h, \underline{\boldsymbol{\rho}}^h, \boldsymbol{\lambda}^h) \in \widetilde{\underline{\boldsymbol{V}}}^h \times \boldsymbol{W}^h \times \underline{\boldsymbol{A}}^h \times \boldsymbol{M}^h$ satisfies* (3.25) *if and only if $\boldsymbol{\lambda}^h$ is the unique solution of*

$$a_h(\boldsymbol{\lambda}^h, \boldsymbol{\mu}) = b_h(\boldsymbol{\mu}), \qquad \text{for all } \boldsymbol{\mu} \in \boldsymbol{M}^h, \qquad (3.27)$$

where the forms are defined by $a_h(\boldsymbol{\mu}, \boldsymbol{\gamma}) := (\underline{\boldsymbol{\sigma}}_{\boldsymbol{\mu}}, \underline{\boldsymbol{\sigma}}_{\boldsymbol{\gamma}})_{\Omega_h}$, $b_h(\boldsymbol{\mu}) := (f, \boldsymbol{u}_{\boldsymbol{\mu}})_{\Omega_h}$, and

$$\underline{\boldsymbol{\sigma}}^h = \underline{\boldsymbol{\sigma}}_{\boldsymbol{\lambda}^h} + \underline{\boldsymbol{\sigma}}_{\boldsymbol{f}}, \qquad \boldsymbol{u}^h = \boldsymbol{u}_{\boldsymbol{\lambda}^h} + \boldsymbol{u}_{\boldsymbol{f}}, \qquad \underline{\boldsymbol{\rho}}^h = \underline{\boldsymbol{\rho}}_{\boldsymbol{\lambda}^h} + \underline{\boldsymbol{\rho}}_{\boldsymbol{f}}.$$

Proof. This proof is substantially similar to a proof in [11], so we will be brief. By linear superposition of the equations (3.26) defining the local solutions, it is obvious that $\underline{\boldsymbol{\sigma}}_{\boldsymbol{\lambda}^h} + \underline{\boldsymbol{\sigma}}_{\boldsymbol{f}}$, $\boldsymbol{u}_{\boldsymbol{\lambda}^h} + \boldsymbol{u}_{\boldsymbol{f}}$, and $\underline{\boldsymbol{\rho}}_{\boldsymbol{\lambda}^h} + \underline{\boldsymbol{\rho}}_{\boldsymbol{f}}$ satisfy the first three equations of the hybridized mixed method (3.25a)–(3.25c). Hence they coincide with $\underline{\boldsymbol{\sigma}}^h$, $\boldsymbol{u}^h$ and $\boldsymbol{\lambda}^h$, respectively, if and only if the last equation (3.25d) is also satisfied. But since (3.26) implies (the details can be worked out along the lines of analogous arguments in the proof of [11, Lemma 2.2])

$$-\langle \boldsymbol{\mu}, \underline{\boldsymbol{\sigma}}_{\boldsymbol{\lambda}^h} \boldsymbol{n} \rangle_{\partial \Omega_h} = (\mathcal{A}\underline{\boldsymbol{\sigma}}_{\boldsymbol{\lambda}^h}, \underline{\boldsymbol{\sigma}}_{\boldsymbol{\mu}})_{\Omega_h},$$

$$-\langle \boldsymbol{\mu}, \underline{\boldsymbol{\sigma}}_{\boldsymbol{f}} \boldsymbol{n} \rangle_{\partial \Omega_h} = (\boldsymbol{f}, \boldsymbol{u}_{\boldsymbol{\mu}})_{\Omega_h},$$

we find that (3.25d) holds if and only if $\boldsymbol{\lambda}^h$ satisfies (3.27). □

4. Error estimate for the stress

We will now establish an *a priori* bound on the distance between $\underline{\boldsymbol{\sigma}}$ and $\underline{\boldsymbol{\sigma}}^h$ in the following norm that depends on the compliance tensor $\mathcal{A}$

$$\|\underline{\boldsymbol{v}}\|_{L^2(\Omega;\mathcal{A})} := \sqrt{(\mathcal{A}\underline{\boldsymbol{v}}, \underline{\boldsymbol{v}})_\Omega}.$$

Since the projection error $\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}$ converges at the same asymptotic rate as the L^2 best approximation error (by Theorem 2.7), the optimal rate of convergence for $\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$ follows from that of $\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$ given below.

Theorem 4.1. *Let $\underline{\boldsymbol{\sigma}}$ be the solution to (1.1) and $\underline{\boldsymbol{\sigma}}^h$ be the approximation given by (1.4). Then,*

$$\|\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|_{L^2(\Omega;\mathcal{A})} \le C(\|\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_{L^2(\Omega;\mathcal{A})} + \|\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega;\mathcal{D})}).$$

where $\underline{\boldsymbol{P}}$ is the L^2 orthogonal projection onto $\underline{\boldsymbol{A}}^h$ and $\mathcal{D}$ is the inverse of $\mathcal{A}$.

Proof. Let us make some preliminary observations on the function to be bounded, namely $\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$. First, from (2.21) we see that div $\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} = \boldsymbol{P}\boldsymbol{f}$, so

$$\text{div}\,(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h) = 0, \tag{4.28}$$

where we used (1.4b). Furthermore, by (1.7b), (1.7c), and (1.4c)

$$(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h, \underline{\boldsymbol{\eta}})_\Omega = 0, \qquad \text{for all } \underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}^h, \tag{4.29}$$

where we have also used the fact the exact solution $\underline{\boldsymbol{\sigma}}$ is symmetric.

These, together with the equations of the method, are enough to prove the error estimate, as we now see. From (1.1b) we obtain

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

which, after introducing $\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\sigma}}$ and $\underline{\boldsymbol{P}}\underline{\boldsymbol{\rho}}$ and rearranging, gives

$$(\mathcal{A}\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_\Omega + (\boldsymbol{u}, \text{div}\,\underline{\boldsymbol{v}})_\Omega + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}}, \underline{\boldsymbol{v}})_\Omega = (\mathcal{A}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}), \underline{\boldsymbol{v}})_\Omega + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}, \underline{\boldsymbol{v}})_\Omega,$$

for all $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}^h$. Subtracting (1.4a) from the above equation, we get that

$$(\mathcal{A}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h), \underline{\boldsymbol{v}})_\Omega + (\boldsymbol{u} - \boldsymbol{u}^h, \text{div}\,\underline{\boldsymbol{v}})_\Omega + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{v}})_\Omega$$
$$= (\mathcal{A}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}), \underline{\boldsymbol{v}})_\Omega + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}, \underline{\boldsymbol{v}})_\Omega,$$

for all $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}^h$. Choosing $\underline{\boldsymbol{v}} = \underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h$ and using (4.28) and (4.29), we obtain

$$\|\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h\|^2_{L^2(\Omega;\mathcal{A})} = (\mathcal{A}(\underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}), \underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}, \underline{\boldsymbol{\Pi}}\,\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}^h)_\Omega.$$

Hence, the proof is finished by applying the Cauchy-Schwarz inequality. □

5. Error estimate for the displacement

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

$$\text{div}\,\underline{\boldsymbol{\psi}} = \boldsymbol{\theta} \quad \text{in } \Omega, \tag{5.30a}$$
$$\mathcal{A}\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{\phi}) = 0 \quad \text{in } \Omega, \tag{5.30b}$$
$$\boldsymbol{\phi} = 0 \quad \text{on } \partial\Omega. \tag{5.30c}$$

We assume that $\underline{\boldsymbol{\psi}} \in \text{dom}\,\underline{\boldsymbol{\Pi}}$. We also assume that

$$\|\underline{\boldsymbol{\psi}}\|_{H^s(\Omega)} + \|\boldsymbol{\phi}\|_{H^{1+s}(\Omega)} \le C\|\boldsymbol{\theta}\|_{L^2(\Omega)}, \tag{5.31}$$

for some $s \geq 0$. Such inequalities are known to hold in several cases, e.g., in the case of planar elasticity with scalar coefficients on a convex domain, its holds [7] with $s = 1$. The main result of this section is the following error estimate.

Theorem 5.1. *If the regularity estimate* (5.31) *holds, then we have*
$$\|\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)} \leq C\, h^s (\|\boldsymbol{\underline{\Pi}}\,\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{L^2(\Omega)} + \|\boldsymbol{\underline{\rho}} - \boldsymbol{P}\,\boldsymbol{\underline{\rho}}\|_{L^2(\Omega)})$$
where $h = \max(h_K : K \in \Omega_h)$.

The proof of this theorem is based on the identity of the following lemma.

Lemma 5.2. *If* $\boldsymbol{\underline{\psi}}$ *is in* $\mathrm{dom}(\boldsymbol{\underline{\Pi}})$ *and* $\boldsymbol{\underline{\xi}}$ *is the antisymmetric part of* $\mathrm{grad}\,\boldsymbol{\phi}$, *then*
$$(\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{\theta})_\Omega = (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}})_\Omega + (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{\underline{\xi}} - \boldsymbol{P}\,\boldsymbol{\underline{\xi}})_\Omega$$
$$- (\boldsymbol{\sigma} - \boldsymbol{\underline{\Pi}}\,\boldsymbol{\sigma}, \mathrm{grad}\,(\boldsymbol{\phi} - \boldsymbol{P}\boldsymbol{\phi}))_\Omega + (\boldsymbol{\underline{\rho}} - \boldsymbol{P}\,\boldsymbol{\underline{\rho}}, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega,$$
where we define $\mathrm{grad}\,(\boldsymbol{P}\boldsymbol{\phi})$ *on each element by* $\mathrm{grad}\,(\boldsymbol{P}\boldsymbol{\phi})|_K := \mathrm{grad}\,((\boldsymbol{P}\boldsymbol{\phi})|_K)$ *for all* $K \in \Omega_h$.

Proof. We begin by splitting the left hand side into two parts T_1 and T_2 as follows:
$$(\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{\theta})_\Omega = (\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}^h, \mathrm{div}\,\boldsymbol{\underline{\psi}})_\Omega \qquad \text{by (5.30a)},$$
$$= (\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}^h, \mathrm{div}\,\boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}})_\Omega \qquad \text{by (2.21) of Theorem 2.7},$$
$$= (\boldsymbol{u} - \boldsymbol{u}^h, \mathrm{div}\,\boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}})_\Omega$$
$$= T_1 + T_2$$

with
$$T_1 := -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}})_\Omega \quad \text{and} \quad T_2 := (\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}})_\Omega.$$
In the last step above, we have used (1.4a) and (1.1b).

Let us first manipulate T_2.
$$T_2 = (\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\psi}})_\Omega + (\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega$$
$$= (\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\psi}})_\Omega + (\boldsymbol{P}\,\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega + (\boldsymbol{\underline{\rho}} - \boldsymbol{P}\,\boldsymbol{\underline{\rho}}, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega$$
$$= (\boldsymbol{\underline{\rho}} - \boldsymbol{P}\,\boldsymbol{\underline{\rho}}, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega.$$
Here we used that $(\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\psi}})_\Omega = 0$ which follows from the fact that $\boldsymbol{\underline{\psi}}$ is symmetric and $\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h$ is antisymmetric. Moreover, we used (1.7b) and (1.7c) to conclude that $(\boldsymbol{P}\,\boldsymbol{\underline{\rho}} - \boldsymbol{\underline{\rho}}^h, \boldsymbol{\underline{\Pi}}\,\boldsymbol{\underline{\psi}} - \boldsymbol{\underline{\psi}})_\Omega = 0$.

For T_1, let us first analyze the following similar term:
$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\underline{\psi}})_\Omega = (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \mathcal{D}\boldsymbol{\underline{\epsilon}}(\boldsymbol{\phi}))_\Omega, \qquad \text{by (5.30b)}$$
$$= (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{\underline{\epsilon}}(\boldsymbol{\phi}))_\Omega, \qquad \text{as } \mathcal{A} \text{ is symmetric.}$$
Letting $\boldsymbol{\underline{\xi}}$ be the antisymmetric part of $\mathrm{grad}\,\boldsymbol{\phi}$ and continuing, we have
$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\underline{\psi}})_\Omega = (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \mathrm{grad}\,\boldsymbol{\phi})_\Omega - (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{\underline{\xi}})_\Omega$$
$$= (\boldsymbol{\sigma} - \boldsymbol{\underline{\Pi}}\,\boldsymbol{\sigma}, \mathrm{grad}\,\boldsymbol{\phi})_\Omega + (\boldsymbol{\underline{\Pi}}\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \mathrm{grad}\,\boldsymbol{\phi})_\Omega$$
$$- (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{\underline{\xi}} - \boldsymbol{P}\,\boldsymbol{\underline{\xi}})_\Omega - (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{P}\,\boldsymbol{\underline{\xi}})_\Omega.$$
Now, $(\boldsymbol{\sigma}^h - \boldsymbol{\underline{\Pi}}\,\boldsymbol{\sigma}, \mathrm{grad}\,\boldsymbol{\phi})_\Omega = 0$ by integration by parts and (4.28). The symmetry of $\boldsymbol{\sigma}$ and the weak symmetry of $\boldsymbol{\sigma}^h$ given by (1.4c) shows that $(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{P}\,\boldsymbol{\underline{\xi}})_\Omega$ is also zero. Hence,
$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\underline{\psi}})_\Omega = (\boldsymbol{\sigma} - \boldsymbol{\underline{\Pi}}\,\boldsymbol{\sigma}, \mathrm{grad}\,(\boldsymbol{\phi} - \boldsymbol{P}\boldsymbol{\phi}))_\Omega - (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \boldsymbol{\underline{\xi}} - \boldsymbol{P}\,\boldsymbol{\underline{\xi}})_\Omega,$$

where we have used that $(\boldsymbol{\sigma} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\sigma}, \operatorname{grad} \boldsymbol{P}\boldsymbol{\phi})_\Omega = 0$, a consequence of (1.7b).

Hence we can rewrite T_1 as

$$T_1 = (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\psi} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\psi})_\Omega + (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\psi})_\Omega$$
$$= (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\psi} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\psi})_\Omega - (\boldsymbol{\sigma} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\sigma}, \operatorname{grad}(\boldsymbol{\phi} - \boldsymbol{P}\boldsymbol{\phi}))_\Omega + (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h, \underline{\boldsymbol{\xi}} - \underline{\boldsymbol{P}\,\boldsymbol{\xi}})_\Omega.$$

Adding these terms equaling T_1 and T_2, the lemma is proved.   □

Proof of Theorem 5.1. The proof reduces to bounding the terms in Lemma 5.2. By standard approximation estimates and Theorem 2.7,

$$\|\boldsymbol{\psi} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\psi}\|_{L^2(\Omega)} \le Ch^s \|\boldsymbol{\psi}\|_{H^s(\Omega)}$$
$$\|\underline{\boldsymbol{\xi}} - \underline{\boldsymbol{P}\,\boldsymbol{\xi}}\|_{L^2(\Omega)} \le Ch^s \|\boldsymbol{\psi}\|_{H^s(\Omega)} \le Ch^s \|\operatorname{grad} \boldsymbol{\phi}\|_{H^s(\Omega)} \le Ch^s \|\boldsymbol{\phi}\|_{H^{1+s}(\Omega)}$$
$$\|\operatorname{grad}(\boldsymbol{\phi} - \boldsymbol{P}\boldsymbol{\phi})\|_{L^2(\Omega)} \le Ch^s \|\boldsymbol{\phi}\|_{H^{1+s}(\Omega)}$$

Now, applying Cauchy-Schwarz inequality to the right hand side terms of the identity of Lemma 5.2, and using the boundedness of $\mathcal{A}$, we obtain

$$(\boldsymbol{P}\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{\theta})_\Omega \le C\,h^s (\|\underline{\boldsymbol{\Pi}}\,\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{L^2(\Omega)} + \|\boldsymbol{P}\,\boldsymbol{\rho} - \boldsymbol{\rho}\|_{L^2(\Omega)}) \|\boldsymbol{\theta}\|_{L^2(\Omega)},$$

where we have also used Theorem 4.1 and the regularity assumption (5.31). The result now follows by taking the supremum over $\boldsymbol{\theta} \in \boldsymbol{L}^2(\Omega)$ of unit norm.   □

6. Error Estimate for the rotation

Theorem 6.1. *We have* $\|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)} \le C(\|\boldsymbol{\sigma} - \underline{\boldsymbol{\Pi}}\,\boldsymbol{\sigma}\|_{L^2(\Omega)} + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{P}\,\boldsymbol{\rho}}\|_{L^2(\Omega)})$.

Proof. Using (1.4a) and (1.1b) we have

$$(\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{v}}) = -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \underline{\boldsymbol{v}}) - (\boldsymbol{u} - \boldsymbol{u}^h, \operatorname{div}\underline{\boldsymbol{v}}) + (\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \boldsymbol{\rho}, \underline{\boldsymbol{v}}),$$

for all $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}^h$. Choose $\underline{\boldsymbol{v}}$ such that $\operatorname{div}\underline{\boldsymbol{v}} = 0$ and $(\underline{\boldsymbol{v}}, \underline{\boldsymbol{\eta}}) = (\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{\eta}})$ for all $\underline{\boldsymbol{\eta}}$ in $\underline{\boldsymbol{A}}^h$. This is possible by Lemma 2.8. Then,

$$\|\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)}^2 = -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \underline{\boldsymbol{v}}) + (\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \boldsymbol{\rho}, \underline{\boldsymbol{v}}),$$
$$\le C(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\|_{L^2(\Omega)} + \|\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}\|_{L^2(\Omega)}) \|\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)},$$

where we have used $\|\underline{\boldsymbol{v}}\|_{L^2(\Omega)} \le C\|\underline{\boldsymbol{P}\,\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h\|_{L^2(\Omega)}$, an estimate given by Lemma 2.8. The result now follows after we apply the triangle inequality and Theorem 4.1.   □

We remark that Lemma 2.8 seems to be indispensable above, whereas the error estimates for $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ did not use Lemma 2.8. Nonetheless, note that one can prove an optimal error estimate for $\underline{\boldsymbol{\rho}}$ without using Lemma 2.8 if one applies Theorem 5.1 assuming full $H^2(\Omega)$-regularity of (5.30) and quasi-uniform meshes.

7. Error estimates for the Lagrange multiplier

The error in the Lagrange multiplier $\boldsymbol{\lambda}^h$ is the subject of this section. We can bound the error by a local argument once we use the hybridized formulation (3.25). Specifically, we use the first equation of the discrete method (3.25a) and the corresponding exact equation (1.1b), which together imply that on any mesh element K, we have

$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \underline{\boldsymbol{v}})_K + (\boldsymbol{u} - \boldsymbol{u}^h, \operatorname{div}\underline{\boldsymbol{v}})_K + (\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}^h, \underline{\boldsymbol{v}})_K + \langle \boldsymbol{u} - \boldsymbol{\lambda}^h, \underline{\boldsymbol{v}}\boldsymbol{n} \rangle_K = 0, \quad (7.32)$$

for all $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}^k(K)$. Now, let $\boldsymbol{P}_\partial$ denote the L^2 orthogonal projection into $\boldsymbol{M}_h$, i.e.,

$$\langle \boldsymbol{P}_\partial\,\boldsymbol{u}, \boldsymbol{\mu} \rangle_F = \langle \boldsymbol{u}, \boldsymbol{\mu} \rangle_F, \qquad \text{for all } \boldsymbol{\mu} \in \boldsymbol{M}_h, \qquad\qquad (7.33)$$

for any face F of any mesh element. Then, since $\operatorname{div}\underline{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^k(K)$ and $\underline{\boldsymbol{v}}\boldsymbol{n} \in \boldsymbol{\mathcal{P}}^k(F)$ for any $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{V}}^k(K)$, we can rewrite (7.32) as

$$(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}}^h),\underline{\boldsymbol{v}})_K+(\boldsymbol{P}\boldsymbol{u}-\boldsymbol{u}^h,\operatorname{div}\underline{\boldsymbol{v}})_K+(\underline{\boldsymbol{\rho}}-\underline{\boldsymbol{\rho}}^h,\underline{\boldsymbol{v}})_K+\langle\boldsymbol{P}_\partial\,\boldsymbol{u}-\boldsymbol{\lambda}^h,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_K = 0. \quad (7.34)$$

The following theorem is proved using (7.34).

Theorem 7.1. *Let $\boldsymbol{P}_\partial$ be the L^2 projection onto $\boldsymbol{M}^h$ defined in (7.33). Then,*

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

where

$$\|\boldsymbol{\mu}\|_h := \sum_{K\in\Omega_h} h_K\|\boldsymbol{\mu}\|^2_{L^2(\partial K)}.$$

Proof. We use (7.34), and we choose $\underline{\boldsymbol{v}}$ there to be the unique function in $\underline{\boldsymbol{V}}^k(K)$ satisfying (2.18) with $\boldsymbol{E} = \boldsymbol{P}_\partial\,\boldsymbol{u}-\boldsymbol{\lambda}^h$ and $\underline{\boldsymbol{G}} = \underline{\boldsymbol{H}} = 0$. Then,

$$\|\boldsymbol{P}_\partial\,\boldsymbol{u}-\boldsymbol{\lambda}^h\|^2_{L^2(\partial K)} = -(\mathcal{A}(\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}}^h),\underline{\boldsymbol{v}})_K - (\boldsymbol{P}\boldsymbol{u}-\boldsymbol{u}^h,\operatorname{div}\underline{\boldsymbol{v}})_K + (\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}}-\underline{\boldsymbol{\rho}},\underline{\boldsymbol{v}})_K.$$

By Corollary 2.6, we know that $\|\underline{\boldsymbol{v}}\|^2_{L^2(K)} \leq Ch_K\|\boldsymbol{P}_\partial\,\boldsymbol{u}-\boldsymbol{\lambda}^h\|^2_{L^2(\partial K)}$. Hence, using the Cauchy-Schwarz inequality and an inverse estimate, we obtain

$$Ch_K\|\boldsymbol{P}_\partial\,\boldsymbol{u}-\boldsymbol{\lambda}^h\|^2_{L^2(\partial K)} \leq h_K^2(\|\underline{\boldsymbol{\sigma}}-\underline{\boldsymbol{\sigma}}^h\|^2_{L^2(K)}+\|\underline{\boldsymbol{\rho}}-\underline{\boldsymbol{P}}\,\underline{\boldsymbol{\rho}}\|^2_{L^2(K)})+\|\boldsymbol{P}\boldsymbol{u}-\boldsymbol{u}^h\|^2_{L^2(K)}.$$

The result follows by summing over $K \in \Omega_h$ and applying Theorems 4.1 and 5.1. □

Note that if the assumptions of Theorem 5.1 hold, its conclusion, combined with Theorem 7.1, yields a superconvergence result for $\boldsymbol{\lambda}^h$.

8. POSTPROCESSING

After computing the mixed approximations $\underline{\boldsymbol{\sigma}}^h$ and $\boldsymbol{u}^h$, it is possible to post-process them to yield a new displacement approximation $\boldsymbol{u}^{h,\star}$ of enhanced accuracy. The postprocessing scheme is local (and hence inexpensive) and is defined following [18]. On each simplex $K \in \Omega_h$, the postprocessed approximation $\boldsymbol{u}^{h,\star}$ is the (unique) function in $\boldsymbol{\mathcal{P}}^{k+1}(K)$ satisfying

$$(\operatorname{grad}\boldsymbol{u}^{h,\star},\operatorname{grad}\boldsymbol{\omega})_K = (\mathcal{A}\underline{\boldsymbol{\sigma}}^h + \underline{\boldsymbol{\rho}}^h,\operatorname{grad}\boldsymbol{\omega})_K \quad \forall\boldsymbol{\omega} \in \tilde{\boldsymbol{\mathcal{P}}}^{k+1}(K),$$

$$(\boldsymbol{u}^{h,\star},\boldsymbol{w})_K = (\boldsymbol{u}^h,\boldsymbol{w})_K \qquad \forall\boldsymbol{w} \in \boldsymbol{\mathcal{P}}^k(K).$$

Then, we have the following result. We omit its proof as it proceeds along the same lines as a proof in Stenberg [18] (except that we use Theorem 5.1 instead).

Theorem 8.1. *Suppose $\boldsymbol{u}$ is in $\boldsymbol{H}^{k+2}(\Omega)$ and the regularity assumption (5.31) holds with $s = 1$. Then*

$$\|\boldsymbol{u}-\boldsymbol{u}^{h,\star}\|_{L^2(\Omega)} \leq C\,h^{k+2}\,|\boldsymbol{u}|_{H^{k+2}(\Omega)}.$$

As a final remark, observe that if we use the hybridized scheme and consequently have access to $\boldsymbol{\lambda}^h$, then we can simplify the postprocessing further. Indeed, setting $\boldsymbol{\zeta}^{h,\star} := \boldsymbol{u}^{h,\star} - \boldsymbol{u}^h$, we have that $\boldsymbol{\zeta}^{h,\star}|_K$ is the unique function in $\boldsymbol{\mathcal{P}}^{k+1}(K)$ satisfying

$$(\operatorname{grad}\boldsymbol{\zeta}^{h,\star},\operatorname{grad}\boldsymbol{\omega})_K = \langle\boldsymbol{u}^h - \boldsymbol{\lambda}^h,(\operatorname{grad}\boldsymbol{\omega})\boldsymbol{n}\rangle_{\partial K} \quad \forall\boldsymbol{\omega} \in \tilde{\boldsymbol{\mathcal{P}}}^{k+1}(K),$$

$$(\boldsymbol{\zeta}^{h,\star},\boldsymbol{\omega})_K = 0 \qquad \forall\boldsymbol{\omega} \in \boldsymbol{\mathcal{P}}^k(K).$$

Here we have used (3.25a) to obtain the first equation. Its advantage is that its right hand side can be computed using quadratures in one less dimension.

References

[1] M. Amara and J.M. Thomas, *Equilibrium finite elements for the linear elastic problem*, Numer. Math. 33 (1979), no. 4, 367–383.

[2] S. Adams and B. Cockburn, *A mixed finite element method for elasticity in three dimensions*, J. Sci. Comput., 25 (2005), pp. 515–521.

[3] D. N. Arnold, G. Awanou, and R. Winther, *Finite elements for symmetric tensors in three dimensions*, Math. Comp., 77 (2008), pp. 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 R. Winther, *Mixed finite elements for elasticity*, Numer. Math., 92 (2002), pp. 401–419.

[7] 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.

[8] D. Boffi, F. Brezzi, and M. Fortin, *Reduced symmetry elements in linear elasticity*, Commun. Pure Appl. Anal., 8 (2009), pp. 95–121.

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

[10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, Amsterdam, 1978.

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

[12] 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.

[13] J. Gopalakrishnan, L. E. García-Castillo, and L. F. Demkowicz, *Nédélec spaces in affine coordinates*, Comput. Math. Appl., 49 (2005), pp. 1285–1294.

[14] P. Monk, *Finite element methods for Maxwell's equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2003.

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

[16] J.-C. Nédélec, *Mixed Finite Elements in $\mathbb{R}^3$*, Numer. Math., 35 (1980), pp. 315–341.

[17] 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.

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

School of Mathematics, University of Minnesota, Minneapolis, MN 55455
E-mail address: cockburn@math.umn.edu

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