2012

A Locking-Free hp DPG Method for Linear Elasticity with Symmetric Stresses

Jamie Bramwell
University of Texas at Austin

Leszek Demkowicz
University of Texas at Austin

Jay Gopalakrishnan
Portland State University, gjay@pdx.edu

Weifeng Qiu
University of Minnesota - Twin Cities

A LOCKING-FREE hp DPG METHOD FOR LINEAR ELASTICITY WITH SYMMETRIC STRESSES

J. BRAMWELL, L. DEMKOWICZ, J. GOPALAKRISHNAN, AND W. QIU

ABSTRACT. We present two new methods for linear elasticity that simultaneously yield stress and displacement approximations of optimal accuracy in both the mesh size h and polynomial degree p. This is achieved within the recently developed discontinuous Petrov-Galerkin (DPG) framework. In this framework, both the stress and the displacement approximations are discontinuous across element interfaces. We study locking-free convergence properties and the interrelationships between the two DPG methods.

1. INTRODUCTION

In this paper we propose a new method for numerically solving the system of equations describing linear elasticity. The accurate computation of stresses is of critical importance in many applications. Yet, many traditional methods only yield approximations to the displacement. This means that stress approximations must be recovered afterward by numerical differentiation. There are newer methods, of the mixed and discontinuous Galerkin (DG) category, which do give direct stress approximations. However, their stability properties, as a function of both h (the mesh size) and p (the polynomial degree of solution approximations) are presently unknown. In this contribution, we bring to the table, a method of the novel discontinuous Petrov-Galerkin (DPG) type, which exhibits stability independent of h and p. The new method is the first hp-optimal method for linear elasticity that can simultaneously approximate the stress and the displacement. We are also able to show, theoretically and practically, that the convergence of the discrete solution does not deteriorate as the Poisson ratio approaches 0.5, i.e., the method does not lock.

To understand how the new DPG method sidesteps the traditional difficulties, let us review the usual difficulties in designing schemes that give direct stress approximations. Mixed methods [2, 5] based on the Hellinger-Reissner variational principle face the problem of designing an approximation space for stresses consisting of matrix functions that are pointwise symmetric. While this is not difficult in itself, when combined with two other additional requirements, difficulties arise. The first requirement is that forces on a mesh face shared by two mesh elements must be in equilibrium. The second requirement is that the conforming stress space, together with a space for displacement approximations, form a stable pair. To put this in mathematical terms, let $\mathbb{M}$ denote the space of real $N \times N$ matrices and let $\mathbb{S}$ denote its subspace of symmetric matrices. The above mentioned stress properties imply that the exact stress σ on an elastic body occupying $\Omega \subseteq \mathbb{R}^N$ lies in the

2000 *Mathematics Subject Classification.* 65N30, 65L12.

Key words and phrases. linear elasticity, hp method, stress symmetry, discontinuous Galerkin, Petrov-Galerkin, DPG method, ultraweak formulation.

Corresponding author: Weifeng Qiu (qiuxa001@ima.umn.edu).

This work was supported in part by the NSF under grant DMS-1211635. The authors also gratefully acknowledge the support from the IMA (Minneapolis) during their 2010-11 program.

space

$$H(\mathrm{div}, \Omega; \mathbb{S}) = \{\sigma \in L^2(\Omega; \mathbb{S}) : \mathrm{div}\, \sigma \in L^2(\Omega, \mathbb{R}^N)\}. \tag{1.1}$$

(Here, the set of functions from Ω into $\mathbb{X}$ whose components are square integrable on Ω is denoted by $L^2(\Omega, \mathbb{X})$, for $\mathbb{X} = \mathbb{S}, \mathbb{M}, \mathbb{R}^N$ etc.) Mixed methods must use conforming finite element subspaces of $H(\mathrm{div}, \Omega; \mathbb{S})$. Although such spaces are known [2], they have too many unknowns (e.g., their lowest order space has 162 degrees of freedom on a single tetrahedral element). Such rich spaces seem to be necessary to satisfy both the first (conformity) and the second (discrete stability) requirement. In contrast, our new DPG methods change the game by separating the approximation and the stability issues.

The DPG method uses a *weaker* variational formulation for the same problem. In this formulation, σ is sought in the space $L^2(\Omega, \mathbb{M})$, in contrast to the space $H(\mathrm{div}, \Omega; \mathbb{S})$ above. Since $L^2(\Omega, \mathbb{M})$, has no interelement continuity constraints, we are able to design an approximating finite element subspace trivially. Furthermore, due to the nonstandard stabilization mechanism of the DPG scheme, the discrete stress space can be chosen to be a subspace of $L^2(\Omega, \mathbb{S})$, i.e., the method gives stresses that are exactly (point-wise) symmetric. One can equally well choose stress approximations in a subspace of $H(\mathrm{div}, \Omega; \mathbb{S})$, disregarding discrete stability considerations. The stability of the DPG method is inherited from the well-posedness of the new ultraweak formulation. Of course, it is by no means trivial to prove this well-posedness (and most of the analysis in this paper is devoted to it). It is provable by adopting a Petrov-Galerkin framework where trial and test spaces are different. For any given trial space, we can locally obtain a test space that is guaranteed to yield stability.

Test spaces that guarantee stability can be obtained by following the DPG methodology introduced in [14, 15]. Our initial idea was as follows: If one uses DG spaces, then given any test space norm, one can *locally* construct test spaces that yield solutions that are the best approximations in a dual norm on the trial space. This dual norm was called the "energy norm". However, we realized [16] that these energy norms are often complicated to work with once we move beyond one-dimensional problems. But we turned the tables in [13, 34], by showing that given a desirable norm in which one wants the DPG solutions to converge, there is a way to calculate the matching test space norm. We refer to this norm as the "optimal norm" on the test space (see § 5.2). The catch is that the optimal norm is nonlocal. Its use would make the computation of a basis for the test space too expensive. Hence, we have been in pursuit of norm equivalences. If one uses, in place of the optimal norm, an equivalent, but localizable test norm, then the DPG method, instead of delivering the very best approximation, delivers a quasioptimal approximation, i.e., the discretization error is bounded by a scalar multiple of the best approximation error. This approach was applied to a one-dimensional wave propagation problem in [34]. A number of further theoretical tools were needed to develop an error analysis for multidimensional problems. These, in the context of the simple Poisson equation, appear in [13]. In this paper, we further generalize these tools to the case of the elasticity problem and introduce new tools to prove locking-free estimates.

Before we proceed to the details of this DPG method, let us mention several alternative solutions to handle the difficulty of constructing approximating subspaces of (1.1). One approach is to relax the symmetry constraint of stresses by using a Lagrange multiplier. This means that σ is sought in

$$H(\mathrm{div}, \Omega; \mathbb{M}) = \{\sigma \in L^2(\Omega; \mathbb{M}) : \mathrm{div}\, \sigma \in L^2(\Omega, \mathbb{R}^N)\}$$

(cf. (1.1)), a space for which finite elements are easier to design. This avenue gave rise to mixed methods with weakly imposed symmetry [4, 9, 22, 29, 30, 32]. Yet another avenue is to keep the stress symmetry, but relax the $H(\text{div})$-conformity. This yielded non-conforming methods, e.g., [6, 23, 27]. However, none of these methods have been proved to be *hp*-optimal. The closest attempt to an *hp*-method is [30] which studies a variable degree mixed method, but does not show how the error estimates depend on p. In contrast, we will prove that the (two) DPG methods we present in this paper are *hp*-optimal. Furthermore, since the DPG method can be reinterpreted as a least squares method in a nonstandard inner product, it yields matrix systems that are symmetric and positive definite (despite having both stress and displacement as unknowns).

As mentioned above, there are two new DPG formulations in this paper. The first is easy to derive and is a natural extension of our work on the Poisson equation in [13]. The second differs from the first due to the presence of a scaled Lagrange multiplier. The multiplier serves to obtain the extra stability required to prove the locking-free convergence estimates.

In the next section, we introduce the first DPG method. We also state the main convergence result for the first method. The second method and its convergence result is presented in Section 3. Then, in Section 4, we study the relationship between these two methods, discovering when they are equivalent. The proofs of the above mentioned two convergence theorems appear in Section 5. As corollaries to these convergence theorems, we obtain h and p convergence rates in Section 6. In Appendix A, we present a result on a mixed method that we crucially use in our proofs.

2. The first DPG method

In this section, we present the derivation of the first of our two DPG methods for linear elasticity. We also state a convergence theorem, which will be proved in a later section.

Linear elasticity is described by two equations. The first is the constitutive equation

$$A\sigma = \varepsilon(u) \tag{2.1a}$$

and the second is the equilibrium equation

$$\operatorname{div}\sigma = f. \tag{2.1b}$$

These equations are imposed on a domain $\Omega \subseteq \mathbb{R}^N$ and the space dimension N equals 2 or 3. We assume that Ω is a bounded open subset of $\mathbb{R}^N$ with connected Lipschitz boundary. The stress $\sigma(x)$ is a function taking values in $\mathbb{S}$ and its divergence ($\operatorname{div}\sigma$) is taken row-wise. The strain tensor is denoted by $\varepsilon(u) = (\operatorname{grad} u + (\operatorname{grad} u)')/2 = \operatorname{sym}\operatorname{grad} u$ where the prime ($'$) denotes matrix transpose, and $\operatorname{sym} M = (M + M')/2$. The material properties are incorporated through the compliance tensor $A(x)$ in (2.1a) which at each $x \in \Omega$, is a fourth order tensor mapping $\mathbb{S}$ into $\mathbb{S}$. The vector function $u : \Omega \mapsto \mathbb{R}^N$ denotes the displacement field engendered by the body force $f : \Omega \mapsto \mathbb{R}^N$. We consider the simple boundary condition

$$u = 0 \qquad \text{on } \partial\Omega \tag{2.1c}$$

which signifies that the elastic body is clamped on the boundary $\partial\Omega$. (We will remark on extending the method to other boundary condition later – see Remark 3.1.)

To motivate the derivation of the first scheme, we multiply the equations of (2.1) by test functions $\tau : \Omega \mapsto \mathbb{S}$ and $v : \Omega \mapsto \mathbb{R}^N$, supported on a domain K. We temporarily assume τ

and v to be smooth so we can integrate by parts to get

$$(A\sigma, \tau)_K + (u, \operatorname{div} \tau)_K - \langle u, \tau\, n\rangle_{1/2,\partial K} = 0, \tag{2.2a}$$

$$(\sigma, \nabla v)_K - \langle v, \sigma\, n\rangle_{1/2,\partial K} = (f, v)_K. \tag{2.2b}$$

Here n denotes the outward unit normal on the boundary of the domain under consideration, $(\cdot,\cdot)_K$ denotes the integral over K of an appropriate inner product (Frobenius, dot product, or scalar multiplication) of its arguments, and $\langle \cdot, l\rangle_{1/2,\partial K}$ denotes the action of a functional $l \in H^{-1/2}(\partial K)$. Here and throughout, we use standard Sobolev spaces without explanation, e.g., $H^1(\Omega, \mathbb{R}^N) = \{v \in L^2(\Omega, \mathbb{R}^N) : \operatorname{grad} v \in L^2(\Omega, \mathbb{M})\}$.

Now, assume that the domain Ω where the boundary value problem (2.1) is posed, admits a disjoint partitioning into open "elements" K, i.e., $\cup\{\bar{K} :\ K \in \Omega_h\} = \bar{\Omega}$. We need not assume that elements are of any particular shape, only that the mesh elements $K \in \Omega_h$ have Lipschitz and *piecewise planar* boundaries, i.e., in two space dimensions, K is a Lipschitz polygon, and in three space dimensions, K is a Lipschitz polyhedron. We now sum up the equations of (2.2) element by element to obtain

$$(A\sigma, \tau)_{\Omega_h} + (u, \operatorname{div} \tau)_{\Omega_h} - \langle u, \tau\, n\rangle_{\partial\Omega_h} = 0 \tag{2.3a}$$

$$(\sigma, \nabla v)_{\Omega_h} - \langle v, \sigma\, n\rangle_{\partial\Omega_h} = (f, v)_{\Omega_h} \tag{2.3b}$$

$$(\sigma, q)_{\Omega_h} = 0. \tag{2.3c}$$

Here, we have additionally imposed the symmetry of the stress tensor by the last equation (where q is a skew-symmetric matrix valued test function on Ω) and used the following notations:

$$(r, s)_{\Omega_h} = \sum_{K\in\Omega_h} (r, s)_K, \quad \langle w, l\rangle_{\partial\Omega_h} = \sum_{K\in\Omega_h} \langle w, l\rangle_{1/2,\partial K}. \tag{2.4}$$

The notation $\partial\Omega_h$ is used for the collection $\{\partial K :\ K \in \Omega_h\}$. Note that it will be clear from the context if differential operators are calculated element by element or globally, e.g., div in (2.3) is calculated piecewise, while in (1.1) it is the global.

The equations of (2.3) motivate the following rigorous functional framework for an ultraweak variational formulation. We let the traces of u and σ in the terms $\langle u, \tau\, n\rangle_{\partial\Omega_h}$ and $\langle v, \sigma\, n\rangle_{\partial\Omega_h}$ to be new unknowns, which we call the *numerical trace* and the *numerical flux*, resp. The ultraweak variational formulation seeks $(\sigma, u, \hat{u}, \hat{\sigma}_n)$ in the *trial space*

$$U = L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times H_0^{1/2}(\partial\Omega_h; \mathbb{V}) \times H^{-1/2}(\partial\Omega_h; \mathbb{V}), \tag{2.5a}$$

satisfying

$$b\big(\, (\sigma, u, \hat{u}, \hat{\sigma}_n),\ (\tau, v, q)\, \big) = l(\tau, v, q) \qquad \forall\, (\tau, v, q) \in V, \tag{2.5b}$$

where the *test space* V is defined by

$$V = H(\operatorname{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V}) \times L^2(\Omega_h; \mathbb{K}), \tag{2.5c}$$

where $\mathbb{K}$ denotes the subspace of $\mathbb{M}$ consisting of all skew-symmetric matrices, and the bilinear form $b(\cdot,\cdot)$ and the linear form $l(\cdot)$ are motivated by (2.3). Namely,

$$\begin{aligned} b((\sigma, u, \hat{u}, \hat{\sigma}_n), (\tau, v, q)) &= (A\sigma, \tau)_{\Omega_h} + (u, \operatorname{div} \tau)_{\Omega_h} - \langle \hat{u}, \tau\, n\rangle_{\partial\Omega_h} \\ &\quad + (\sigma, \nabla v)_{\Omega_h} - \langle v, \hat{\sigma}_n\rangle_{\partial\Omega_h} + (\sigma, q)_{\Omega_h}, \\ l(\tau, v, q) &= (f, v). \end{aligned} \tag{2.5d}$$

In the notations of (2.5a) and (2.5c), $\mathbb{V}$ denotes the vector space $\mathbb{R}^N$ and

$$\begin{aligned}
H_0^{1/2}(\partial\Omega_h;\mathbb{V}) &= \{\eta : \exists w \in H_0^1(\Omega;\mathbb{V}) \text{ with } \eta|_{\partial K} = w|_{\partial K}, \forall K \in \Omega_h\},\\
H^{-1/2}(\partial\Omega_h;\mathbb{V}) &= \{\eta \in \prod_{K\in\Omega_h} H^{-1/2}(\partial K;\mathbb{V}) : \exists q \in H(\operatorname{div},\Omega;\mathbb{M}) \text{ with}\\
&\qquad\qquad \eta|_{\partial K} = q\, n|_{\partial K}, \forall K \in \Omega_h\},\\
H(\operatorname{div},\Omega_h;\mathbb{S}) &= \{\tau : \tau|_K \in H(\operatorname{div},K;\mathbb{S}), \forall K \in \Omega_h\},\\
H^1(\Omega_h;\mathbb{V}) &= \{v : v|_K \in H^1(K;\mathbb{V}), \forall K \in \Omega_h\}.
\end{aligned}$$

The norms on $H_0^{1/2}(\partial\Omega_h;\mathbb{V})$ and $H^{-1/2}(\partial\Omega_h;\mathbb{V})$ are defined by

$$\|\hat u\|_{H_0^{1/2}(\partial\Omega_h)} = \inf\{\|w\|_{H^1(\Omega)} : \forall w \in H_0^1(\Omega;\mathbb{V}) \text{ with } \hat u - w|_{\partial K} = 0\}, \tag{2.6}$$

$$\|\hat\sigma_n\|_{H^{-1/2}(\partial\Omega_h)} = \inf\{\|q\|_{H(\operatorname{div},\Omega)} : \forall q \in H(\operatorname{div},\Omega;\mathbb{M}) \text{ with } \hat\sigma_n - q\, n|_{\partial K} = 0\}. \tag{2.7}$$

The trial and test norms are defined by

$$\|(\sigma,u,\hat u,\hat\sigma_n)\|_U^2 = \|\sigma\|_\Omega^2 + \|u\|_\Omega^2 + \|\hat u\|^2_{H_0^{1/2}(\partial\Omega_h)} + \|\hat\sigma_n\|^2_{H^{-1/2}(\partial\Omega_h)}, \tag{2.8a}$$

$$\|(\tau,v,q)\|_V^2 = \|\tau\|^2_{H(\operatorname{div},\Omega_h)} + \|v\|^2_{H^1(\Omega_h)} + \|q\|^2_\Omega. \tag{2.8b}$$

Here the norms on $H^1(\Omega_h;\mathbb{V})$ and $H(\operatorname{div},\Omega_h;\mathbb{S})$ are defined by

$$\|v\|^2_{H^1(\Omega_h)} = (v,v)_{\Omega_h} + (\operatorname{grad} v, \operatorname{grad} v)_{\Omega_h}, \tag{2.9}$$

$$\|\tau\|^2_{H(\operatorname{div},\Omega_h)} = (\tau,\tau)_{\Omega_h} + (\operatorname{div}\tau, \operatorname{div}\tau)_{\Omega_h}. \tag{2.10}$$

This completes the description of our new infinite dimensional DPG variational formulation.

Remark 2.1. Like in many other numerical formulations, in the DPG formulation (2.5), we need to the extend the domain of the compliance tensor $A(x)$ from $\mathbb{S}$ to $\mathbb{M}$. There are many ways to perform this extension. To choose one, decompose $\mathbb{M}$ orthogonally (in the Frobenius inner product) into $\mathbb{K}$ and $\mathbb{S}$. A standard way to extend $A(x)$ from $\mathbb{S}$ to $\mathbb{M}$ is to define $A(x)\kappa = \kappa$ for all κ in $\mathbb{K}$. Then, whenever the original $A(x)$ is self-adjoint and positive definite on $\mathbb{S}$, the extended $A(x)$ is also self-adjoint and positive definite on $\mathbb{M}$.

We assume throughout that $A(x)$ (i.e., its above mentioned extension) is self-adjoint and positive definite uniformly on Ω. We also assume that the components of (the extended) A are in $L^\infty(\Omega)$.

Next, we describe the discrete DPG scheme. This is done following verbatim the abstract setup in [13, Section 2] (see also [15]). Accordingly, we define the *trial-to-test* operator $T : U \mapsto V$ by

$$(T\mathcal{U},\mathcal{V})_V = b(\mathcal{U},\mathcal{V}), \qquad \forall\, \mathcal{V} \in V \text{ and } \forall \mathcal{U} \in U. \tag{2.11}$$

We select *any* finite dimensional subspace $U_h \subseteq U$ and set the corresponding finite dimensional test space by

$$V_h = T(U_h).$$

Then the DPG approximation $(\sigma_h, u_h, \hat u_h, \hat\sigma_{n,h}) \in U_h$ satisfies

$$b\big(\,(\sigma_h,u_h,\hat u_h,\hat\sigma_{n,h}),\ (\tau,v,q)\,\big) = l(\tau,v,q) \qquad \forall\,(\tau,v,q) \in V_h. \tag{2.12}$$

The distance between this approximation and the exact solution can be bounded as stated in the next theorem.

Theorem 2.2 (Quasioptimality)**.** *Let $U_h \subseteq U$. Then,* (2.5b) *has a unique solution* $(\sigma, u, \hat{u}, \hat{\sigma}_n) \in U$ *and* (2.12) *has a unique solution* $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}) \in U_h$. *Moreover, there is a* $C^{(1)} > 0$ *independent of the subspace* U_h *and the partition* Ω_h *such that*

$$\mathcal{D} \;\le\; C^{(1)}\,\mathcal{A},$$

where $\mathcal{D}$ *is the discretization error and* $\mathcal{A}$ *is the error in best approximation by* U_h*, defined by*

$$\mathcal{D} = \|\sigma - \sigma_h\|_{L^2(\Omega)} + \|u - u_h\|_{L^2(\Omega)} + \|\hat{u} - \hat{u}_h\|_{H_0^{1/2}(\partial\Omega_h)} + \|\hat{\sigma}_n - \hat{\sigma}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)}.$$

$$\mathcal{A} = \inf_{(\rho_h, w_h, \hat{z}_h, \hat{\eta}_{n,h}) \in U_h} \Big(\|\sigma - \rho_h\|_{L^2(\Omega)} + \|u - w_h\|_{L^2(\Omega)} + \|\hat{u} - \hat{z}_h\|_{H_0^{1/2}(\partial\Omega_h)} + \|\hat{\sigma}_n - \hat{\eta}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)} \Big).$$

This result is comparable to Cea's lemma in traditional finite element theory. Of importance is the independence of $C^{(1)}$ with respect to U_h. Specifically, we are interested in setting U_h to hp-finite element subspaces with extreme variations in h and p to capture singularities or thin layers in solutions. In this case, the constant $C^{(1)}$, being independent of U_h, is independent of *both* the mesh size h and the polynomial degree p. As such, this forms the first method for linear elasticity with provably hp-optimal convergence rates of the same order for σ and u. Although several mixed methods yielding good approximations to σ are known, proving their hp-optimality requires proving an inf-sup condition carefully tracking the dependence of constants on p, a feat yet to be achieved. For a proof of Theorem 2.2, see Section 5.

3. The second DPG method

The robustness of numerical methods with respect to the Poisson ratio is an important consideration in computational mechanics. Methods that are not robust exhibit *locking*. Note that we did not assert in Theorem 2.2 that the constant $C^{(1)}$ is independent of the Poisson ratio. However, in all our numerical experiments (see Section 7), the method showed locking-free convergence. This section serves as a first step towards explaining this locking-free behavior theoretically.

The second DPG method given below is designed so that we can establish locking-free convergence with respect to the Poisson ratio. It has one more trial variable. In this section, we will assert its locking-free convergence properties, restricting ourselves to isotropic materials. In the next section, we will provide a sufficient condition under which the first and the second DPG methods are equivalent. This gives theoretical insight into the locking-free behavior of *both* the first and the second methods.

Let us begin by defining the essential infimum

$$Q_0 = \operatorname*{ess\,inf}_{x\in\Omega} \big(\, \mathrm{tr}(A(x)I) \,\big). \tag{3.1}$$

Obviously, $Q_0 > 0$ for an isotropic material with Poisson ratio $\nu < 0.5$. The second method is motivated by the same integration by parts as in (2.3), but with the following additional observation in mind: If we set $\tau = I$ in (2.3) and recall that $u = 0|_{\partial\Omega}$, then we have that

$$\int_\Omega \mathrm{tr}\, A\sigma = 0. \tag{3.2}$$

Imposing this condition via a Lagrange multiplier, we obtain another ultraweak formulation with the following bilinear and linear forms:

$$\begin{aligned} b^{(2)}((\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, q, \beta)) = {} & (A\sigma, \tau)_{\Omega_h} + (u, \operatorname{div} \tau)_{\Omega_h} - \langle \hat{u}, \tau\, n\rangle_{\partial\Omega_h} + Q_0^{-1}(\alpha I, A\tau)_{\Omega_h} \\ & + (\sigma, \nabla v)_{\Omega_h} + (\sigma, q)_{\Omega_h} - \langle v, \hat{\sigma}_n\rangle_{\partial\Omega_h} \\ & + Q_0^{-1}(A\sigma, \beta I)_{\Omega_h} \qquad (3.3a) \\ l^{(2)}(\tau, v, q, \beta) = {} & (f, v). \end{aligned}$$

Here I is the identity matrix. The trial space is now set to

$$U^{(2)} = L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times H_0^{1/2}(\partial\Omega_h; \mathbb{V}) \times H^{-1/2}(\partial\Omega_h; \mathbb{V}) \times \mathbb{R} \qquad (3.3b)$$

and the test space is set to

$$V^{(2)} = H(\operatorname{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V}) \times L^2(\Omega_h; \mathbb{K}) \times \mathbb{R}. \qquad (3.3c)$$

By (3.2), the solution (σ, u) of (2.1) will be the solution of (3.3) with

$$\hat{u} = u|_{\partial\Omega_h}, \quad \hat{\sigma}_n = \sigma n|_{\partial\Omega_h}, \quad \text{and} \quad \alpha = 0.$$

However, more work is needed to conclude similar statements at the discrete level (see the next section).

Remark 3.1. In the case of mixed boundary conditions we must add the term

$$-Q_0^{-1}\langle \hat{u}, (\beta I)n\rangle_{\partial\Omega_h} \qquad (3.4)$$

to the the expression (3.3a). Note that this term vanishes in the case of kinematic boundary conditions analyzed in this paper. However, for more general boundary conditions, $\hat{u}$ can be be nonzero on the parts of the boundary where traction conditions are imposed. Hence (3.4) simplifies to a boundary integral that is nonzero in general.

The second DPG method is obtained by constructing a discrete scheme as before from the ultraweak formulation (following [13, Section 2]). The trial-to-test operator in this case (cf. (3.5)) is $T^{(2)} : U^{(2)} \mapsto V^{(2)}$ by

$$(T^{(2)}\mathcal{U}, \mathcal{V})_{V^{(2)}} = b^{(2)}(\mathcal{U}, \mathcal{V}), \qquad \forall\, \mathcal{V} \in V^{(2)}. \qquad (3.5)$$

Let $U_h^{(2)} \subseteq U^{(2)}$ be any finite dimensional subspace. We set $V_h^{(2)} = T^{(2)}(U^{(2)})$. The second DPG approximation $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, \alpha_h) \in U_h^{(2)}$ satisfies

$$b^{(2)}\big(\,(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, \alpha_h),\ (\tau, v, q, \beta)\,\big) = l((\tau, v, q, \beta)) \qquad \forall\,(\tau, v, q, \beta) \in V_h^{(2)}. \qquad (3.6)$$

As in the case of the first DPG method, we are able to prove a quasioptimality result (see the next theorem) bounding the discretization error

$$\begin{aligned} \mathcal{D}^{(2)} = {} & \|\sigma - \sigma_h\|_{L^2(\Omega)} + \|u - u_h\|_{L^2(\Omega)} + |\alpha - \alpha_h| \\ & + \|\hat{u} - \hat{u}_h\|_{H_0^{1/2}(\partial\Omega_h)} + \|\hat{\sigma}_n - \hat{\sigma}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)}. \end{aligned}$$

by the error in best approximation

$$\begin{aligned} \mathcal{A}^{(2)} = \inf_{(\rho_h, w_h, \hat{z}_h, \hat{\eta}_{n,h}, \gamma_h) \in U_h} \Bigg(& \|\sigma - \rho_h\|_{L^2(\Omega)} + \|u - w_h\|_{L^2(\Omega)} + |\alpha - \gamma_h| \\ & + \|\hat{u} - \hat{z}_h\|_{H_0^{1/2}(\partial\Omega_h)} + \|\hat{\sigma}_n - \hat{\eta}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)} \Bigg). \end{aligned}$$

However, we are also able to prove a stronger result under the following assumption.

Assumption 3.1 (Isotropic material). We assume that

$$A\tau = P\tau_D + Q\frac{\mathrm{tr}(\tau)}{N}I \tag{3.7}$$

where

$$\tau_D = \tau - \frac{\mathrm{tr}(\tau)}{N}I,$$

for any τ in $\mathbb{M}$, and P and Q are positive scalar functions on Ω. (Then, obviously $Q \geq Q_0$ for the Q_0 defined in (3.1).) When (3.7) holds, we also define

$$B = Q_0^{-1}\|Q\|_{L^\infty(\Omega)}, \tag{3.8}$$

$$P_0 = \operatorname*{ess\,inf}_{x\in\Omega} P(x). \tag{3.9}$$

(Note that (3.7) is assumed to hold for all $\tau \in \mathbb{M}$.)

Remark 3.2. When we consider isotropic materials, we do not extend A from $\mathbb{S}$ to $\mathbb{M}$ in the way suggested in Remark 2.1. Instead, we assume that it is extended from $\mathbb{S}$ to $\mathbb{M}$ by $A\kappa = P\kappa$ for all $\kappa \in \mathbb{K}$. This ensures that (3.7) holds for all $\tau \in \mathbb{M}$, not just for all $\tau \in \mathbb{S}$.

Theorem 3.3 (Quasioptimality of the second DPG method)**.** *Let* $U_h^{(2)} \subseteq U^{(2)}$*. Then,* (3.3) *has a unique solution* $(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in U^{(2)}$ *and* (3.6) *has a unique solution* $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, \alpha_h) \in U_h^{(2)}$*. Moreover, there is a* $C^{(2)} > 0$ *independent of the subspace* $U_h^{(2)}$ *and the partition* Ω_h *such that*

$$\mathcal{D}^{(2)} \;\leq\; C^{(2)}\,\mathcal{A}^{(2)}.$$

If in addition, Assumption 3.1 holds, then $C^{(2)}$ *can be chosen to be*

$$C^{(2)} = \bar{c}P_0^{-1}(\|A\| + B)^3 B^4 (\|A\| + P_0 + 1)^2, \tag{3.10}$$

a constant independent of Q_0*, and consequently the method does not lock. (Here, the positive constant* $\bar{c}$ *is independent of* A*.)*

The proof of this theorem appears in Section 5.

4. The relationship between the two DPG methods

In this section we will establish that for homogeneous isotropic materials the two DPG methods are equivalent. We will also show that despite the additional trial variable, the second DPG method can be solved in essentially the same cost as the first.

4.1. The equivalence.

Recall that the first DPG variational formulation uses the trial space

$$U = L^2(\Omega;\mathbb{M}) \times L^2(\Omega;\mathbb{V}) \times H_0^{1/2}(\partial\Omega_h;\mathbb{V}) \times H^{-1/2}(\partial\Omega_h;\mathbb{V}),$$

while the second uses $U \times \mathbb{R}$. We begin with the following simple lemma.

Lemma 4.1. *For any* $\mathcal{U} \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n)$ *in* U*, let*

$$(\tau, v, q) = T\mathcal{U}.$$

Then, with $\mathcal{U}^{(2)} \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n, 0)$*,*

$$T^{(2)}\mathcal{U}^{(2)} = (\tau, v, q, \beta),$$

with $\beta = Q_0^{-1}(A\sigma, I)_\Omega$*.*

Proof. By the definition of T, we have, for any $(\delta_\tau, \delta_v, \delta_q) \in V = H(\mathrm{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V}) \times L^2(\Omega_h; \mathbb{K})$,

$$\begin{aligned}\big((\tau, v, q),\ (\delta_\tau, \delta_v, \delta_q) \big)_V &= (\tau, \delta_\tau)_\Omega + (\mathrm{div}\,\tau, \delta_\tau)_{\Omega_h} + (v, \delta_v)_\Omega + (\nabla v, \nabla \delta_v)_{\Omega_h} + (q, \delta_q)_\Omega \\ &= (A\sigma, \delta_\tau)_\Omega + (u, \mathrm{div}\,\delta_\tau)_{\Omega_h} - \langle \hat{u}, \delta_\tau n\rangle_{\partial\Omega_h} + (\sigma, \nabla\delta_v)_{\Omega_h} - \langle \delta_v, \hat{\sigma}_n\rangle_{\partial\Omega_h} + (\sigma, \delta_q)_\Omega.\end{aligned}$$

Therefore choosing

$$\beta = Q_0^{-1}(A\sigma, I)_\Omega,$$

we obviously obtain

$$\begin{aligned}&\big((\tau, v, q, \beta), (\delta_\tau, \delta_v, \delta_q, \delta_\beta) \big)_{V^{(2)}} \\ &\qquad = (\tau, \delta_\tau)_\Omega + (\mathrm{div}\,\tau, \delta_\tau)_{\Omega_h} + (v, \delta_v)_\Omega + (\nabla v, \nabla\delta_v)_{\Omega_h} + (q, \delta_q)_\Omega + \beta\delta_\beta \\ &\qquad = (A\sigma, \delta_\tau)_\Omega + (u, \mathrm{div}\,\delta_\tau)_{\Omega_h} - \langle \hat{u}, \delta_\tau n\rangle_{\partial\Omega_h} + (\sigma, \nabla\delta_v)_{\Omega_h} - \langle \delta_v, \hat{\sigma}_n\rangle_{\partial\Omega_h} \\ &\qquad + (\sigma, \delta_q)_\Omega + Q_0^{-1}(A\sigma, \delta_\beta I)_\Omega \\ &\qquad = b^{(2)}\big((\sigma, u, \hat{u}, \hat{\sigma}_n, 0),\ (\tau, v, q, \beta)\big)\end{aligned}$$

for any $(\delta_\tau, \delta_v, \delta_q, \delta_\beta) \in V \times \mathbb{R} = V^{(2)}$. This finishes the proof. □

We use the above result together with the following assumption to prove the equivalence. The assumption essentially states that the material is homogeneous and isotropic and that the discrete trial space at least contains two specific functions related to the identity matrix.

Assumption 4.1. Suppose Assumption 3.1 (on isotropy) holds. Let the discrete trial subspace $U_h \subseteq U$ of the first DPG method be

$$U_h = \Sigma_h \times W_h \times M_h \times F_h, \tag{4.1}$$

and let the second DPG method use the trial space $U_h \times \mathbb{R}$. We assume that

(1) $Q(x) = Q_0$ for all $x \in \Omega$,
(2) $I \in \Sigma_h$,
(3) $In|_{\partial\Omega_h} \in F_h$.

Lemma 4.2. *If Assumption 4.1 holds, then* $(I, 0, 0) \in V_h \equiv T(U_h)$.

Proof. By virtue of the assumption, the trial function $\mathcal{U} = (\sigma, u, \hat{u}, \hat{\sigma}_n)$ with $\sigma = I, u = 0, \hat{\sigma}_n = In|_{\partial\Omega_h}$, and $\hat{u} = 0$, is in U_h. Hence, $(\tau, v, q) \equiv T\mathcal{U}$ is in V_h and satisfies

$$\begin{aligned}(\tau, \delta_\tau)_\Omega + (\mathrm{div}\,\tau, \mathrm{div}\,\delta_\tau)_{\Omega_h} + (v, \delta_v)_\Omega + (\nabla v, \nabla\delta_v)_{\Omega_h} + (q, \delta_q)_\Omega \\ = (AI, \delta_\tau)_\Omega + (I, \nabla\delta_v)_{\Omega_h} - \langle \delta_v, In\rangle_{\partial\Omega_h} + (I, \delta_q)_\Omega\end{aligned} \tag{4.2}$$

for all $(\delta_\tau, \delta_v, \delta_q) \in V = H(\mathrm{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V}) \times L^2(\Omega_h; \mathbb{K})$. The last term in (4.2) vanishes due to the skew symmetry of δ_q. Integration by parts shows that $(I, \nabla\delta_v)_{\Omega_h} - \langle \delta_v, In\rangle_{\partial\Omega_h} = 0$ also. Hence, we conclude that $(\tau, v, q) = (Q_0 I, 0, 0)$ is the unique solution of (4.2). Since (τ, v, q) is in V_h, we have proved the lemma. □

Lemma 4.3. *If Assumption 4.1 holds, then*

$$(I, 0, 0, 0) = T^{(2)}(0, 0, 0, 0, 1).$$

Proof. Let $(\tau, v, q, \beta) = T^{(2)}(0, 0, 0, 0, \alpha)$. Then by the definition of $T^{(2)}$, we have

$$\begin{aligned}(\tau, \delta_\tau)_\Omega + (\mathrm{div}\,\tau, \mathrm{div}\,\delta_\tau)_{\Omega_h} + (v, \delta_v)_\Omega + (\nabla v, \nabla\delta_v)_{\Omega_h} + (q, \delta_q)_\Omega + \beta\delta_\beta \\ = Q_0^{-1}(\alpha I, A\delta_\tau)_\Omega\end{aligned}$$

for all $(\delta_\tau, \delta_v, \delta_q, \delta_\beta) \in V^{(2)} = H(\mathrm{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V}) \times L^2(\Omega_h; \mathbb{K}) \times \mathbb{R}$. Putting $\alpha = 1$, and using the symmetry of A, the right hand side $Q_0^{-1}(\alpha I, A\delta_\tau)_\Omega = Q_0^{-1}(AI, \delta_\tau)_\Omega$, which due to the assumption on A equals $(I, \delta_\tau)_\Omega$, i.e.,

$$(\tau, \delta_\tau)_\Omega + (\mathrm{div}\,\tau, \mathrm{div}\,\delta_\tau)_{\Omega_h} = (I, \delta_\tau)_\Omega.$$

It is now obvious that $(\tau, v, q, \beta) = (I, 0, 0, 0)$. □

Theorem 4.4. *Suppose Assumption 4.1 holds. Then* $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h})$ *solves the first DPG method* (2.12) *if and only if* $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, 0)$ *is the discrete solution of the second DPG method* (3.6).

Proof. Suppose $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h})$ solves the first DPG method (2.5). To show that the function $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, 0)$ satisfies the second DPG method (3.3), we have to show that the equations

$$(A\sigma_h, \tau)_{\Omega_h} + (u_h, \mathrm{div}\,\tau)_{\Omega_h} - \langle \hat{u}, \tau\, n\rangle_{\partial\Omega_h} + Q_0^{-1}(\alpha I, A\tau)_{\Omega_h} = 0 \tag{4.3a}$$

$$(\sigma_h, \nabla v)_{\Omega_h} - \langle v, \hat{\sigma}_{n,h}\rangle_{\partial\Omega_h} = (f, v) \tag{4.3b}$$

$$(\sigma_h, q)_{\Omega_h} = 0 \tag{4.3c}$$

$$Q_0^{-1}(A\sigma, \beta I)_{\Omega_h} = 0, \tag{4.3d}$$

hold, with $\alpha = 0$, for all $(\tau, v, q, \beta) \in T^{(2)}(U_h \times \mathbb{R})$. To this end, we proceed in two steps.

First, we consider test functions of the type $T^{(2)}(U_h \times \{0\})$. By virtue of Lemma 4.1, these test functions are in $T(U_h) \times \mathbb{R}$. Hence by the fact that $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h})$ solves the equation of the first DPG method (2.12) for all $(\tau, v, q) \in T(U_h)$, we observe that (4.3a), (4.3b) and (4.3c) hold. To show that (4.3d) also holds, we observe that because of Lemma 4.2, we may put $(\tau, v, q) = (\beta I, 0, 0)$, for any $\beta \in \mathbb{R}$, in (2.12) to get

$$(A\sigma_h, \beta I)_{\Omega_h} + (u_h, \mathrm{div}\,\beta I)_{\Omega_h} - \langle \hat{u}_h, \beta I n\rangle_{\partial\Omega_h} = 0.$$

Multiplying by Q_0^{-1} and simplifying, we obtain (4.3d) for any β.

Second, consider test functions of the type $T^{(2)}(\{(0,0,0,0)\} \times \mathbb{R})$. But by Lemma 4.3, we know that $T^{(2)}(0,0,0,0,1) = (I, 0, 0, 0)$. So to show that (4.3) holds for this test function, it suffices to prove that (4.3a) holds with $\tau = I$. But this follows from Lemma 4.2, which shows that $(I, 0, 0)$ is in V_h.

Conversely, suppose $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h}, 0)$ satisfies the second DPG method (3.6). Then, $(\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h})$ must satisfy the first DPG method, for if not, there must be another function $(\sigma'_h, u'_h, \hat{u}'_h, \hat{\sigma}'_{n,h})$ solving the first DPG method. But then, the already proved implication shows that $(\sigma'_h, u'_h, \hat{u}'_h, \hat{\sigma}'_{n,h}, 0)$ must solve the second DPG method. This contradicts the unique solvability of the second DPG method asserted in Theorem 3.3. □

4.2. **Solving the second DPG system.** The relationship between the two methods revealed above, can be utilized to obtain a useful strategy for solving the second DPG system.

To understand the linear systems that result from both methods, we assume that a basis for U_h is made of local (standard finite element) functions

$$e_i = (\sigma_i, u_i, \hat{u}_i, \hat{\sigma}_{n,i})$$

for $i = 1, 2, \ldots, m$. Then the corresponding basis for V_h is furnished by $t_j = Te_i$. Hence the $m \times m$ stiffness matrix of the first DPG method E has entries

$$E_{ij} = b(e_j, t_i).$$

It is easily seen that this matrix is symmetric and positive definite. (This is a general property of DPG stiffness matrices – see [13] or [15].) In addition, E is sparse due to the locality of the basis functions.

Now, consider the stiffness matrix of the second DPG method. Here, we need a basis for $U_h \times \mathbb{R}$. It is natural to take as basis for $U_h \times \mathbb{R}$, the $m+1$ functions defined by

$$e_i^{(2)} = \begin{cases} (e_i, 0), & \text{for all } i = 1, 2, \ldots, m, \\ (0,0,0,0,1), & \text{for } i = m+1. \end{cases}$$

But we should then note that the stiffness matrix

$$E_{ij}^{(2)} = b^{(2)}(e_j^{(2)}, T^{(2)} e_i^{(2)})$$

is no longer sparse. This is because $T^{(2)} e_i^{(2)}$ is not locally supported. Indeed, if we write $T^{(2)} e_i^{(2)}$ as $(\tau_i, v_i, q_i, \beta_i)$, then β_i can be globally supported even if $e_i^{(2)}$ is local. Thus, although $E^{(2)}$ is symmetric and positive definite, one may be led into concluding that the second DPG method is too expensive due to the non-sparsity.

However, this is not the case. Below, we will show how to solve a system $E^{(2)} x^{(2)} = y^{(2)}$ by solving a system $Ex = y$ and performing a few additional inexpensive steps. A key observation is that $\tilde{E}$ is a rank-one perturbation of E.

Proposition 4.1. *Decompose the matrix $E^{(2)}$ into*

$$E^{(2)} = \begin{bmatrix} \tilde{E} & c \\ c' & d \end{bmatrix},$$

where $c \in \mathbb{R}^m$. Then

$$\tilde{E} = E + \ell \ell'$$

where $\ell \in \mathbb{R}^m$ is defined by $\ell_j = Q_0^{-1}(A\sigma_j, I)_\Omega$.

Proof. Let $(\tau_i, v_i, q_i) = T e_i$. Then, by Lemma 4.1, $T^{(2)} e_i^{(2)} = (\tau_i, v_i, q_i, \beta_i)$ with

$$\beta_i = Q_0^{-1}(A\sigma_i, I)_\Omega.$$

Hence,

$$\begin{aligned} \tilde{E}_{ij} &= b^{(2)}\big(e_j^{(2)}, T^{(2)} e_i^{(2)}\big) \\ &= b^{(2)}\big(e_j^{(2)}, (\tau_i, v_i, q_i, \beta_i)\big) \\ &= b^{(2)}\big(e_j^{(2)}, (\tau_i, v_i, q_i, 0)\big) + b^{(2)}\big(e_j^{(2)}, (0,0,0,\beta_i)\big) \\ &= b(e_j, T e_i) + Q_0^{-1}(\beta_i I, A\sigma_j)_\Omega \\ &= E_{ij} + \beta_i \beta_j. \end{aligned}$$

The result follows because $\beta_i = \ell_i$. □

A consequence of this proposition is that we can invert $\tilde{E}$ using the Sherman-Morrison formula [25], namely

$$(E + \ell\ell')^{-1} = E^{-1} - a(E^{-1}\ell)(E^{-1}\ell)', \quad \text{with} \quad a = \frac{1}{1 + \ell' E^{-1} \ell}. \tag{4.4}$$

Therefore, to conclude this discussion, consider the matrix system arising from the second DPG method (3.6):

$$\begin{bmatrix} \tilde{E} & c \\ c' & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} g \\ 0 \end{bmatrix}.$$

Since $\tilde{E}$ is nonsingular, the solution is given by

$$y = -(d - c'\tilde{E}^{-1}c)^{-1}c'\tilde{E}^{-1}g, \qquad x = \tilde{E}^{-1}g - \tilde{E}^{-1}cy.$$

Hence the solution of the second DPG method can be obtained by solving the two linear systems $\tilde{E}x_c = c$ and $\tilde{E}x_g = g$. Each of these systems can be solved using formula (4.4) at essentially the same cost as solving a system involving E.

5. Error analysis

In this section, we prove Theorems 2.2 and 3.3. We will give the proof in full detail for the more difficult case of Theorem 3.3 first. Afterward, we will indicate the minor modification required to prove Theorem 2.2 in a similar fashion. The plan is to use the abstract DPG framework developed in [13, 14, 15, 16, 34], summarized next.

5.1. **Abstract quasioptimality.** Let X with norm $\|\cdot\|_X$ be a reflexive Banach space over $\mathbb{R}$, Y with norm $\|\cdot\|_Y$ be a Hilbert space over $\mathbb{R}$ with inner product $(\cdot,\cdot)_Y$, and $b(\cdot,\cdot) : X \times Y \mapsto \mathbb{R}$ be a bilinear form. The abstract trial-to-test operator $T : X \mapsto Y$ is defined – as before – by $(Tu, v)_Y = b(u, v)$ for all $u \in X$ and $y \in Y$. We can write the DPG method using an arbitrary subspace $X_h \subseteq X$ even in this generality. Let $Y_h = T(X_h)$.

Theorem 5.1 (see [13, Theorem 2.1])**.** *Suppose $x \in X$ and $x_h \in X_h$ satisfy*

$$\begin{aligned} b(x, y) &= l(y) && \forall\, v \in Y, \\ b(x_h, y) &= l(y) && \forall\, y \in Y_h. \end{aligned}$$

Assume that

$$\{w \in X : \ b(w, v) = 0, \ \forall v \in Y\} = \{0\} \tag{5.1}$$

and that there are positive constants C_1, C_2 such that

$$C_1\|y\|_Y \le \|y\|_{\text{opt},Y} \le C_2\|y\|_Y, \qquad \forall y \in Y, \tag{5.2}$$

where the so-called "optimal norm" is defined by

$$\|y\|_{\text{opt},Y} = \sup_{xw\in X} \frac{b(x, y)}{\|x\|_X}. \tag{5.3}$$

Then

$$\|x - x_h\|_X \le \frac{C_2}{C_1} \inf_{z_h\in X_h} \|x - z_h\|_X.$$

5.2. **The optimal test norms of both methods.** We now see what is the norm defined by (5.3) in the context of the first and second DPG methods. From the structure of the bilinear form it is easy to see that, for the first DPG method, the optimal test norm is

$$\begin{aligned} \|(\tau, v, q)\|_{\mathrm{opt},V}^2 &= \sup_{0 \neq (\sigma, u, \hat{u}, \hat{\sigma}_n) \in U} \frac{b((\sigma, u, \hat{u}, \hat{\sigma}_n), (\tau, v, q))^2}{\|(\sigma, u, \hat{u}, \hat{\sigma}_n)\|_U^2} \\ &= \|A\tau + \nabla v + q\|_{\Omega_h}^2 + \|\operatorname{div}\tau\|_{\Omega_h}^2 + \|[\tau\, n]\|_{\partial\Omega_h}^2 + \|[vn]\|_{\partial\Omega_h}^2, \end{aligned} \tag{5.4}$$

where the "jump" terms are defined by

$$\|[\tau\, n]\|_{\partial\Omega_h} := \sup_{0 \neq u \in H_0^1(\Omega;\mathbb{V})} \frac{\langle u, \tau\, n\rangle_{\partial\Omega_h}}{\|u\|_{H^1(\Omega)}} \tag{5.5a}$$

$$\|[vn]\|_{\partial\Omega_h} := \sup_{0 \neq \sigma \in H(\operatorname{div},\Omega;\mathbb{M})} \frac{\langle v, \sigma\, n\rangle_{\partial\Omega_h}}{\|\sigma\|_{H(\operatorname{div},\Omega)}}. \tag{5.5b}$$

Similarly, for the second DPG method, taking the supremum over its trial space, we have

$$\begin{aligned} \|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}^2 = {} & \|A\tau + \nabla v + q + Q_0^{-1}\beta A I\|_{\Omega_h}^2 + \|\operatorname{div}\tau\|_{\Omega_h}^2 \\ & + \|[\tau\, n]\|_{\partial\Omega_h}^2 + \|[vn]\|_{\partial\Omega_h}^2 + \left| Q_0^{-1} \int_\Omega \operatorname{tr}(A\tau)\right|^2. \end{aligned} \tag{5.6}$$

In either case, the optimal norms are inconvenient for practical computations, due to the last two jump terms. These terms would make the trial-to-test computation in (3.5) non-local. Therefore, a fundamental ingredient in our ensuing analysis is the proof of equivalence of the optimal norm with the simpler "broken" norm (for the first DPG method)

$$\|\, (\tau, v, q)\, \|_V^2 = \|\tau\|_{H(\operatorname{div},\Omega_h)}^2 + \|v\|_{H^1(\Omega_h)}^2 + \|q\|_\Omega^2 \tag{5.7}$$

which does contain the jump terms. For the second DPG method, we will similarly prove that (5.6) is equivalent to following norm

$$\|\, (\tau, v, q, \beta)\, \|_{V^{(2)}}^2 = \|\tau\|_{H(\operatorname{div},\Omega_h)}^2 + \|v\|_{H^1(\Omega_h)}^2 + \|q\|_\Omega^2 + |\beta|^2. \tag{5.8}$$

These equivalences would verify condition (5.2), so we would be in a position to apply Theorem 5.1. Before we proceed to prove these, let us verify the other condition (5.1). We begin with a necessary preliminary.

5.3. **Korn inequalities.** We will need two well-known inequalities due to Korn. The *first* Korn inequality asserts the existence of a constant $C > 0$ such that

$$\|v\|_{H^1(\Omega)} \le C\|\varepsilon(v)\|_\Omega \quad \forall v \in H_0^1(\Omega), \tag{5.9a}$$

while the *second* Korn inequality gives a constant $C > 0$ such that

$$\|v\|_{H^1(\Omega)} \le C\,(\|v\|_\Omega + \|\varepsilon(v)\|_\Omega) \quad \forall v \in H^1(\Omega). \tag{5.9b}$$

Above and in the remainder, we use C to denote a generic constant whose value, although possibly different at different occurrences, will remain independent of the discrete approximation spaces. These inequalities can be found in many references. E.g., for (5.9a), see [26, eq. (2.7)], and for (5.9b), see [28, Section 1.12 in Chapter 6].

5.4. **Uniqueness for the second DPG method.** In this subsection, we verify condition (5.1) for the second DPG method.

Lemma 5.2. *With $U^{(2)}$ and $V^{(2)}$ as set in* (3.3) *suppose* $(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in U^{(2)}$ *satisfies*

$$b^{(2)}((\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, q, \beta)) = 0, \qquad \forall (\tau, v, q, \beta) \in V^{(2)}. \tag{5.10}$$

Then $(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) = 0$.

Proof. Equation (5.10) is the same as

$$(A\sigma, \tau)_K + (u, \operatorname{div} \tau)_K - \langle \hat{u}, \tau\, n\rangle_{\partial K} + Q_0^{-1}(\alpha A I, \tau)_K = 0 \qquad \forall \tau \in H(\operatorname{div}, K; \mathbb{S}) \tag{5.11a}$$

$$(\sigma, \operatorname{grad} v)_K - \langle \hat{\sigma}_n, v\rangle_{\partial K} = 0 \qquad \forall v \in H^1(K; \mathbb{V}) \tag{5.11b}$$

$$\int_\Omega \operatorname{tr}(A\sigma) = 0. \tag{5.11c}$$

Here, $\sigma \in L^2(\Omega_h; \mathbb{S})$ because

$$(\sigma, q)_\Omega = 0 \quad \forall q \in L^2(\Omega_h; \mathbb{K}).$$

We take $\tau \in \mathcal{D}(K; \mathbb{S})$ and $v \in \mathcal{D}(K; \mathbb{V})$ arbitrarily. Here, as usual, $\mathcal{D}(D, \mathbb{X})$ denotes the space of infinitely smooth functions from D into $\mathbb{X}$ that are compactly supported in D. Then we have

$$A\sigma - \varepsilon(u) + Q_0^{-1}\alpha A I = 0 \quad \text{in } K \tag{5.12a}$$

$$\operatorname{div} \sigma = 0 \quad \text{in } K \tag{5.12b}$$

in the sense of distributions, for every $K \in \Omega_h$. In particular, this implies that $u \in H^1(K; \mathbb{V})$ by ((5.12a) and the second Korn inequality (5.9b)) and $\sigma \in H(\operatorname{div}, K; \mathbb{S})$ by (5.12b).

Now we claim that we also have

$$\hat{\sigma}_n|_{\partial K} = \sigma\, n|_{\partial K} \qquad \text{and} \qquad u|_{\partial K} = \hat{u}|_{\partial K}. \tag{5.13}$$

The first identity in (5.13) is obtained by integrating (5.11b) by parts and using (5.12b) to find that $\langle v, \hat{\sigma}_n - \sigma\, n\rangle_{1/2, \partial K} = 0$ for every $v \in H^1(\Omega_h; \mathbb{V})$. Note that this identity implies that $\sigma \in H(\operatorname{div}, \Omega; \mathbb{S})$.

To prove the second identity of (5.13), we need to proceed a bit differently. Let F be an arbitrary face of ∂K, say in three dimensions (the two dimensional case is similar and simpler). We will now show that given any $\eta \equiv (\eta_i) \in \mathcal{D}(F; \mathbb{V})$, there is a $\tau \in H(\operatorname{div}, K; \mathbb{S})$ such that $\tau n \in L^2(\partial K; \mathbb{V})$ is supported only on $F \subseteq \partial K$ such that $\tau n|_F = \eta$. To this end, we may, without loss of generality, assume that F is contained in the xy-plane, so that $n = (0, 0, 1)'$. Then, set

$$\tau = \chi(x, y, z) \begin{bmatrix} 0 & 0 & \eta_1(x, y) \\ 0 & 0 & \eta_2(x, y) \\ \eta_1(x, y) & \eta_2(x, y) & \eta_3(x, y) \end{bmatrix} \tag{5.14}$$

where χ is an infinitely smooth cut-off function such that (i) the support of $\chi(x, y, 0)$ is compactly contained in F and contains the support of η, and (ii) $\chi \equiv 1$ on the support of η. Clearly we can find such a cut-off function, and furthermore, construct it so that τn vanishes on all other faces of K.

We use such τ to prove the second one in (5.13). First observe that from (5.10), we have $(A\sigma, \tau)_K + (u, \operatorname{div} \tau)_K - \langle \hat{u}, \tau\, n\rangle_{\partial K} + Q_0^{-1}(\alpha A I, \tau)_K = 0$, for all $\tau \in H(\operatorname{div}, K; \mathbb{S})$. Integrating by parts (which is permissible since by (5.11), $u \in H^1(K;\mathbb{V})$ and $\sigma \in H(\mathrm{div}, K;\mathbb{S})$), and using (5.11a),

$$\langle u - \hat{u}, \tau n\rangle_{\partial K} = 0, \qquad \forall \tau \in H(\mathrm{div}, K;\mathbb{S}). \tag{5.15}$$

Choosing τ as in (5.14), this implies

$$\int_F (u - \hat{u})\, \eta \; ds = 0, \qquad \forall \eta \in \mathcal{D}(F,\mathbb{V}).$$

Hence $u|_F = \hat{u}|_F$ in $L^2(F)$ and this holds for all faces of ∂K. This proves the second identity of (5.13), which implies that $u \in H^1_0(\Omega;\mathbb{V})$ (after also noting that $\hat{u}|_{\partial\Omega} = 0$).

Next, we choose $\tau = I$ on Ω and sum the terms in (5.11a) over all $K \in \Omega_h$. Using the fact that $u|_{\partial K} = \hat{u}|_{\partial K}$, we have that

$$\int_\Omega \mathrm{tr}(A\sigma) + Q_0^{-1}\alpha \int_\Omega \mathrm{tr}(AI) = 0.$$

The first term vanishes due to (5.11c). Hence we have shown that $\alpha = 0$.

Now, choose $\tau = \sigma$ and $v = u$ in (5.11a)–(5.11b). Summing up these equations and canceling terms after integrating by parts, we find that

$$(A\sigma,\sigma)_{\Omega_h} + \langle u, \sigma\, n\rangle_{\partial\Omega_h} - \langle \hat{u}, \sigma\, n\rangle_{\partial\Omega_h} - \langle u, \hat{\sigma}_n\rangle_{\partial\Omega_h} = 0. \tag{5.16}$$

Since $u \in H^1_0(\Omega;\mathbb{V})$ the last term on the left had side vanishes. Furthermore, since we already showed that the interelement jumps of σn are zero, the penultimate term $\langle \hat{u}, \sigma\, n\rangle_{\partial\Omega_h}$ also vanishes. Note finally that $\langle u, \sigma n\rangle_{\partial\Omega_h} = \langle u, \sigma n\rangle_{\partial\Omega} = 0$ as $u \in H^1_0(\Omega)$. Thus, (5.16) implies that $\sigma = 0$.

Since σ vanishes, by (5.12a), we have that $\varepsilon(u) = 0$. Since $u \in H^1_0(\Omega;\mathbb{V})$, by the first Korn inequality (5.9a) we find that $u = 0$. Since both σ and u vanish, by (5.13), $\hat{u}$ and $\hat{\sigma}_n$ also vanish. □

5.5. **An inf-sup condition.** In this subsection, we verify that the lower bound in the condition (5.2) holds for the second DPG method. The lower bound is the same as an inf-sup condition due to the definition of the optimal norm. To prove this inf-sup condition, we use a modification of the mixed method for linear elasticity with weakly imposed symmetry, given in Appendix A.

Lemma 5.3. *There is a positive constant C_1 such that for any $(\tau, v, q, \beta) \in V^{(2)}$,*

$$C_1 \|(\tau, v, q, \beta)\|_{V^{(2)}} \le \|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}.$$

If, in addition, Assumption 3.1 holds, then the constant C_1 can be chosen to be

$$C_1^{-1} = \bar{c}_2 P_0^{-1}(\|A\| + B)^2 B^4 (\|A\| + P_0 + 1)^2, \tag{5.17}$$

where $\bar{c}_2$ is a positive constant independent of A.

Proof. Given $(\tau, v, q, \beta) \in V^{(2)}$, we solve the mixed method (A.1) with data $F_1 = \tau, F_2 = -v, F_3 = q$ and $F_4 = N^{-1}\beta I/|\Omega|$, to get $(\sigma, u, \rho, a) \in H(\mathrm{div},\Omega;\mathbb{M}) \times L^2(\Omega;\mathbb{V}) \times L^2(\Omega;\mathbb{K}) \times \mathbb{R}$. According to Theorem A.1, the component u is in $H^1_0(\Omega;\mathbb{V})$ and

$$\|\sigma\|_{H(\mathrm{div},\Omega)} + \|u\|_{H^1(\Omega)} + \|\rho\|_\Omega + |a| \le C_0(\|\tau\|_\Omega + \|v\|_\Omega + \|q\|_\Omega + |\beta|). \tag{5.18}$$

Within an element K, we can find the strong form of the equations in (A.1) by choosing infinitely smooth test functions that are compactly supported on K. We obtain

$$A\sigma - \nabla u + \rho + aQ_0^{-1}AI = \tau, \tag{5.19a}$$
$$\operatorname{div}\sigma = -v, \tag{5.19b}$$
$$\operatorname{skw}\sigma = q, \tag{5.19c}$$

together with the last equation (A.1d) which can be restated simply as

$$Q_0^{-1}\int_\Omega \operatorname{tr}(A\sigma) = \beta. \tag{5.19d}$$

These, together with integration by parts, imply that

$$\begin{aligned}
&\|\tau\|_\Omega^2 + \|v\|_\Omega^2 + \|q\|_\Omega^2 + |\beta|^2 \\
&\quad= (A\sigma - \nabla u + \rho + aQ_0^{-1}AI, \tau)_\Omega - (\operatorname{div}\sigma, v)_\Omega + (\operatorname{skw}\sigma, q)_\Omega + \beta Q_0^{-1}\int_\Omega \operatorname{tr}(A\sigma) \\
&\quad= (\sigma, A\tau)_{\Omega_h} + (u, \operatorname{div}\tau)_{\Omega_h} - \langle u, \tau n\rangle_{\partial\Omega_h} + (\rho, \tau)_\Omega + aQ_0^{-1}\int_\Omega \operatorname{tr}(A\tau) \\
&\qquad\quad + (\sigma, \nabla v)_{\Omega_h} - \langle v, \sigma n\rangle_{\partial\Omega_h} + (\operatorname{skw}\sigma, q)_\Omega + (\sigma, \beta Q_0^{-1}AI)_{\Omega_h} \\
&\quad= (\sigma, A\tau)_{\Omega_h} + (u, \operatorname{div}\tau)_{\Omega_h} - \langle u, \tau n\rangle_{\partial\Omega_h} + aQ_0^{-1}\int_\Omega \operatorname{tr}(A\tau) \\
&\qquad\quad + (\sigma, \nabla v)_{\Omega_h} - \langle v, \sigma n\rangle_{\partial\Omega_h} + (\sigma, q)_\Omega + (\sigma, \beta Q_0^{-1}AI)_{\Omega_h}
\end{aligned}$$

since τ is symmetric. Rearranging and applying Cauchy-Schwarz inequality,

$$\begin{aligned}
&\|\tau\|_\Omega^2 + \|v\|_\Omega^2 + \|q\|_\Omega^2 + |\beta|^2 \\
&\quad= (\sigma, A\tau + \nabla v + q + \beta Q_0^{-1}AI)_{\Omega_h} + (u, \operatorname{div}\tau)_{\Omega_h} + aQ_0^{-1}\int_\Omega \operatorname{tr}(A\tau) \\
&\qquad\quad - \langle u, \tau n\rangle_{\partial\Omega_h} - \langle v, \sigma n\rangle_{\partial\Omega_h} \\
&\quad\le \|\sigma\|_\Omega \|A\tau + \nabla v + q + \beta Q_0^{-1}AI\|_{\Omega_h} + \|u\|_\Omega \|\operatorname{div}\tau\|_{\Omega_h} + |a|\cdot|Q_0^{-1}\int_\Omega \operatorname{tr}(A\tau)| \\
&\qquad\quad + \|[\tau\, n]\|_{\partial\Omega_h}\|u\|_{H^1(\Omega)} + \|[vn]\|_{\partial\Omega_h}\|\sigma\|_{H(\operatorname{div},\Omega)} \\
&\quad\le 2\,\|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}} \left(\|\sigma\|_{H(\operatorname{div},\Omega)}^2 + \|u\|_{H^1(\Omega)}^2 + \|\rho\|_\Omega^2 + |a|^2\right)^{1/2}
\end{aligned}$$

By (5.18), we have that

$$\|\tau\|_\Omega^2 + \|v\|_\Omega^2 + \|q\|_\Omega^2 + |\beta|^2 \le 4C_0^2\,\|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}^2. \tag{5.20}$$

Furthermore, since $\|A\tau + \nabla v + q + \beta Q_0^{-1}AI\|_{\Omega_h} \le \|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}$, by triangle inequality,

$$\|\nabla v\|_{\Omega_h} \le \|A\|\|\tau\|_\Omega + \|q\|_\Omega + \|AI\|Q_0^{-1}|\beta| + \|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}.$$

which implies that

$$\|\nabla v\|_{\Omega_h} \le c_1\|(\tau, v, q, \beta)\|_{opt,V^{(2)}}.$$

for a positive constant c_1.

If, in addition the material is isotropic in the sense of Assumption 3.1, then using the notations of the assumption, the constant c_1 can be chosen to be

$$c_1 = 2(\|A\| + B)C_0.$$

Finally, since $\|\operatorname{div}\tau\|_{\Omega_h} \le \|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}}$, we can control all terms in forming the norm $\|\operatorname{div}\tau\|_{\Omega_h}$, i.e.,

$$\|(\tau, v, q, \beta)\|_{V^{(2)}} \le c_2(2C_0 + c_1)\|(\tau, v, q, \beta)\|_{opt,V^{(2)}}$$

with a constant c_2 is independent of A. The lemma follows with $C_1^{-1} = c_2(2C_0 + c_1)$. In the case of isotropic material, observe that

$$\begin{aligned} C_1^{-1} &= c_2(2C_0 + c_1) = 2c_2(\|A\| + B + 1)C_0 \le 4c_2(\|A\| + B)C_0 \\ &\le 4c_2(\|A\| + B)\bar{c}_1 P_0^{-1} B^4(\|A\| + P_0 + 1)^2(\|A\| + B) \\ &\le \bar{c}_2 P_0^{-1}(\|A\| + B)^2 B^4(\|A\| + P_0 + 1)^2, \end{aligned}$$

where $\bar{c}_2$ is a positive constant independent of A. ☐

5.6. **Upper bound.** Now we show that the upper inequality of condition (5.2) can be verified for the second DPG method.

Lemma 5.4. *There is a positive constant C_2 such that for any $(\tau, v, q, \beta) \in V^{(2)}$,*

$$\|(\tau, v, q, \beta)\|_{\mathrm{opt},V^{(2)}} \le \bar{c}_3(\|A\| + B)\|(\tau, v, q, \beta)\|_{V^{(2)}}.$$

Here, $\bar{c}_3$ is a positive constant independent of A.

Proof. Let us first prove an upper bound for the jump terms. Integrating by parts locally and applying Cauchy-Schwarz inequality,

$$\begin{aligned} \big\|[\tau\, n]\big\|_{\partial\Omega_h} &= \sup_{w\in H_0^1(\Omega;\mathbb{V})} \frac{\langle w, \tau\, n\rangle_{\partial\Omega_h}}{\|w\|_{H^1(\Omega)}} = \sup_{w\in H_0^1(\Omega;\mathbb{V})} \frac{(\operatorname{grad} w, \tau)_{\Omega_h} + (w, \operatorname{div}\tau)_{\Omega_h}}{\|w\|_{H^1(\Omega)}} \\ &\le \|\tau\|_{H(\operatorname{div},\Omega_h)}. \end{aligned}$$

We use a similar argument for the other jump, i.e.,

$$\begin{aligned} \big\|[vn]\big\|_{\partial\Omega_h} &= \sup_{\varsigma\in H(\operatorname{div},\Omega;\mathbb{M})} \frac{\langle v, \varsigma\, n\rangle_{\partial\Omega_h}}{\|\varsigma\|_{H(\operatorname{div},\Omega;\mathbb{M})}} = \sup_{\varsigma\in H(\operatorname{div},\Omega;\mathbb{M})} \frac{(\operatorname{grad} v, \varsigma)_{\Omega_h} + (v, \operatorname{div}\varsigma)_{\Omega_h}}{\|\varsigma\|_{H(\operatorname{div},\Omega;\mathbb{M})}} \\ &\le \|v\|_{H^1(\Omega_h)}. \end{aligned}$$

The remainder of the proof is straightforward. ☐

5.7. **Proof of Theorem 2.2.** We apply the abstract result of Theorem 5.1. Assumption (5.1) is verified by Lemma 5.2. The lower inequality of (5.2) is verified by Lemma 5.3, and the upper inequality is verified by Lemma 5.4. ☐

5.8. **Proof of Theorem 3.3.** The analysis of the first DPG method is in many ways simpler than the above detailed analysis of the second DPG method. Just as in the proof of Theorem 2.2, we only need to verify the conditions (5.1) and (5.2) of the abstract result.

The proof of the uniqueness condition (5.1) is similar and simpler than the proof of Lemma 5.2, so we omit it.

The proof of the upper inequality in condition (5.2) is the same the proof of Lemma 5.4.

The proof of the lower inequality in condition (5.2) for the first DPG method is analogous and simpler than the proof of Lemma 5.3. To highlight the main difference, instead of considering the mixed method in Appendix A, we now need only use the standard mixed method with weakly imposed stress symmetry. In other words, the analogue of (5.18) is now obtained as follows: Given $(\tau, v, q) \in V$, we solve the following variational problem to find $(\sigma, u, \rho) \in H(\operatorname{div}, \Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{K})$ satisfying

$$(A\sigma, \delta\tau)_\Omega + (u, \operatorname{div} \delta\tau)_\Omega + (\rho, \delta\tau)_\Omega = (\tau, \delta\tau)_\Omega \qquad \forall \delta\tau \in H(\operatorname{div}, \Omega; \mathbb{M}), \tag{5.21a}$$

$$(\operatorname{div} \sigma, \delta v)_\Omega = -(v, \delta v)_\Omega \qquad \forall \delta v \in L^2(\Omega; \mathbb{V}), \tag{5.21b}$$

$$(\sigma, \delta q)_\Omega = (q, \delta q)_\Omega \qquad \forall \delta q \in L^2(\Omega; \mathbb{K}). \tag{5.21c}$$

Then we use the standard stability estimate [4] for this method to get the analogue of (5.18) and proceed as in the proof of Lemma 5.3. □

6. Examples of trial spaces and convergence rates

The trial subspaces of both the first and the second DPG methods (namely U_h and $U_h^{(2)}$) were unspecified in Theorem 2.2 and theorem 3.3. This section is devoted to two examples of trial spaces and how one can use Theorems 2.2 and 3.3 to predict h and p convergence rates for these examples. The examples we have in mind are DG spaces built on a tetrahedral mesh and a cubic mesh. We only consider the case of the first DPG method (as the same convergence rates can be derived analogously for the second DPG method).

If D is a simplex, let $P_p(D)$ denote the set of functions that are restrictions of (multivariate) polynomials of degree at most p on a domain D. If D is cubic, then we write it as a tensor product of three intervals $D = D_x \otimes D_y \otimes D_z$ and define $Q^{p,q,r}(D) = P_p(D_x) \otimes P_q(D_y) \otimes P_r(D_z)$ which is the space of polynomials of degree at most p, q, r with respect to x, y, z, resp. As with Sobolev spaces, when these notations may also be augmented with a range vector space, i.e., $P_p(D; \mathbb{S})$ denotes the space of symmetric matrix-valued functions whose components are polynomials of degree at most p, etc.

Recall that – see (4.1) – to specify U_h, we must specify its four component spaces. If Ω_h is a tetrahedral mesh, we set

$$\Sigma_{h,p} = \{\rho : \rho|_K \in P_p(K; \mathbb{S})\}, \qquad W_{h,p} = \{v : v|_K \in P_p(K; \mathbb{V})\}, \tag{6.1a}$$

while if Ω_h is a cubic mesh, then we set

$$\Sigma_{h,p} = \{\rho : \rho|_K \in Q^{p,p,p}(K; \mathbb{S})\}, \qquad W_{h,p} = \{v : v|_K \in Q^{p,p,p}(K; \mathbb{V})\}. \tag{6.1b}$$

Note that we have chosen the subspaces to consist of *symmetric* matrix polynomials. This is clearly allowed since the only requirement for the discrete trial subspace was that $U_h \subset U$. (The corresponding automatically generated test space ensures stability of the resulting method. We emphasize that this stabilization mechanism is different from mixed methods.) The numerical flux space is set as follows:

$$F_{h,p} = \{\eta : \ \eta|_E \in P_p(E; \mathbb{V}) \ \forall \text{ mesh faces } E\} \qquad \text{if } \Omega_h \text{ is a tetrahedral mesh}, \tag{6.1c}$$

$$F_{h,p} = \{\eta : \ \eta|_E \in Q^{p,p}(E; \mathbb{V}) \ \forall \text{ mesh faces } E\} \qquad \text{if } \Omega_h \text{ is a cubic mesh}. \tag{6.1d}$$

In either case, we define the numerical trace space by

$$M_{h,p+1} = \{\eta : \exists\, w \in W_{h,p+1} \cap H_0^1(\Omega; \mathbb{V}) \text{ such that } \eta|_{\partial K} = w|_{\partial K} \ \forall K \in \Omega_h\}. \tag{6.1e}$$

Since $p \geq 0$, the space $M_{h,p+1}$ is non-trivial.

Let us apply Theorem 3.3 with these as trial spaces for each solution component. Then, if we know how the best approximation error converges in terms of h and p, we can conclude rates of convergence. It is well known that for $s > 0$,

$$\inf_{w_h \in W_{h,p}} \|u - w_h\|_\Omega \le C h^s p_2^{-s} |u|_{H^s(\Omega)}, \qquad (s \le p+1). \tag{6.2}$$

Here $p_2 = \max(p, 2)$. A similar best approximation estimate obviously holds for σ as well. Note that since the exact stress σ is symmetric, it can be approximated to optimal accuracy by the symmetric subspace $\Sigma_{h,p}$.

Next, consider the flux and trace best approximations in the quotient topology defined by (2.6) and (2.7). Since the exact trace $\hat{u}$ is the trace of the exact solution u, and since the exact flux $\hat{\sigma}_n$ is the trace of the normal components of σ along each interface of $\partial\Omega_h$, we have

$$\inf_{\hat{z}_h \in M_{h,p}} \|\hat{u} - \hat{z}_h\|_{H_0^{1/2}(\partial\Omega_h)} \le \|u - \mathit{\Pi}_{\mathrm{grad}} u\|_{H^1(\Omega)},$$
$$\inf_{\hat{\eta}_{n,h} \in Q_{h,p}} \|\hat{\sigma}_n - \hat{\eta}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)} \le \|\sigma - \mathit{\Pi}_{\mathrm{div}} \sigma\|_{H(\mathrm{div},\Omega)},$$

where $\mathit{\Pi}_{\mathrm{grad}} u \in H_0^1(\Omega; \mathbb{V})$ and $\mathit{\Pi}_{\mathrm{div}} \sigma \in H(\mathrm{div}, \Omega; \mathbb{M})$ are suitable projections, such that their traces $\mathit{\Pi}_{\mathrm{grad}} u|_E$ and $\mathit{\Pi}_{\mathrm{div}} \sigma\, n|_E$ on any mesh face E is in $P_p(E; \mathbb{V})$ or $Q^{p,p}(E; \mathbb{V})$. These conforming projectors providing approximation estimates with constants independent of p are available from recent works in [11, 12, 17, 18, 19]. Specifically, [19, Theorem 8.1] gives

$$\|u - \mathit{\Pi}_{\mathrm{grad}} u\|_{H^1(\Omega)} \le C \ln(p_2)^2\, h^s p_2^{-s} |u|_{H^{s+1}(\Omega)}, \qquad (s \le p), \tag{6.3a}$$
$$\|\sigma - \mathit{\Pi}_{\mathrm{div}} \sigma\|_{L^2(\Omega)} \le C \ln(p_2)\, h^s p_2^{-s} |\sigma|_{H^{s+1}(\Omega)}, \qquad (s \le p+1). \tag{6.3b}$$

whenever $s > 1/2$ for tetrahedral meshes. The same results for cubic meshes are available from [11]. Note that we have chosen $\mathit{\Pi}_{\mathrm{div}}$ to be a projector into the Raviart-Thomas space. The projector into the Raviart-Thomas space satisfies $\mathrm{div}\, \mathit{\Pi}_{\mathrm{div}} \sigma = \mathit{\Pi}_p \,\mathrm{div}\, \sigma$ (where $\mathit{\Pi}_p$ denote the L^2-orthogonal projection into $W_{h,p}$). Hence,

$$\|\,\mathrm{div}(\sigma - \mathit{\Pi}_{\mathrm{div}} \sigma)\|_{L^2(\Omega)} \le C h^s p_2^{-s} |\,\mathrm{div}\, \sigma|_{H^s(\Omega)}, \qquad (s \le p+1). \tag{6.4}$$

Finally, comparing the rates of convergence in (6.2), (6.3) and (6.4), we find that to obtain a full $O(h^{p+1})$ order of convergence, we must increase the polynomial degree of the numerical trace space to $p + 1$. Combining these observations, we have the following corollary.

Corollary 6.1 (h and p convergence rates)**.** *Let Ω_h be a shape regular mesh (either tetrahedral or cubic) and let h denote the maximum of the diameters of its elements. Let $\mathcal{D}$ and $\mathcal{D}^{(2)}$ denote the (previously defined) discretization errors of the first and second DPG methods, resp. Using the spaces defined in* (6.1)*, set*

$$U_h = \Sigma_{h,p} \times W_{h,p} \times M_{h,p+1} \times F_{h,p}$$

for the first DPG method and

$$U_h^{(2)} = \Sigma_{h,p} \times W_{h,p} \times M_{h,p+1} \times F_{h,p} \times \mathbb{R}$$

for the second DPG method. Then

$$\mathcal{D} \le C_I \ln(p_2)^2\, h^s p_2^{-s} (\|\sigma\|_{H^{s+1}(\Omega)} + \|u\|_{H^{s+1}(\Omega)})$$
$$\mathcal{D}^{(2)} \le C_{II} \ln(p_2)^2\, h^s p_2^{-s} (\|\sigma\|_{H^{s+1}(\Omega)} + \|u\|_{H^{s+1}(\Omega)})$$

for all $1/2 < s \leq p+1$. The constants C_I and C_{II} are independent of h and p, but dependent on the shape regularity and A. If Assumption 3.1 (isotropy) holds, then C_{II} is independent of Q_0, so the second estimate does not degenerate as the Poisson ratio goes to 0.5.

In the same way, one can derive convergence rates for other element shapes and spaces from Theorems 2.2 and Theorem 3.3.

Remark 6.1 (Symmetric and conforming stresses). If an application demands stress approximations σ_h that are both symmetric and div-conforming (i.e., if one needs σ_h to be in the space $H(\text{div}, \Omega; \mathbb{S})$ defined in (1.1)), then the DPG method can certainly give such approximations. We only need to choose a trial subspace

$$\Sigma_{h,p} \subset H(\text{div}, \Omega; \mathbb{S}) \tag{6.5}$$

instead of the choice $\Sigma_{h,p} \subset L^2(\Omega, \mathbb{S})$ we made in (6.1) above. Obviously $H(\text{div}, \Omega; \mathbb{S}) \subset L^2(\Omega, \mathbb{M})$, so (6.5), together with the other component spaces as set previously, would result in a trial space U_h that is conforming in our ultraweak variational framework. Hence, Theorems 2.2 and 3.3 continue to apply. Notice that stability of the resulting DPG method is ensured even with this choice because the method adapts its test space to any given trial subspace. The first example of $\Sigma_{h,p}$ satisfying (6.5) that comes to mind is the finite element of [2]. However their space is too rich because they had to ensure a discrete inf-sup condition. Since we have separated out the stability issue, we have other simpler and inexpensive options. E.g., we may choose $\Sigma_{h,p}$ to consist of symmetric matrix functions, each of whose entries are in $L_{h,p} = \{\rho \in H^1(\Omega, \mathbb{S}) : \rho|_K \in P_p(K, \mathbb{S})\}$ (i.e., continuous Lagrange finite element functions). The locality of our test space construction is not destroyed with this choice. Moreover, p-optimal interpolation estimates are known for this space, so we can proceed as above to state an analogue of Corollary 6.1. Of course, the same remarks also apply for displacement approximations, e.g., we may choose H^1-conforming subspaces to approximate the displacement.

7. Numerical results

In this section, we present numerical results for the first DPG method using two test cases: a smooth solution on a square domain and a singular solution on an L-shaped domain. All numerical experiments were conducted using a pre-existing *hp*-adaptive finite element package [10].

7.1. **Discrete spaces.** Following [13], we considered a 2D domain Ω divided into conforming or 1-irregular quadrilateral meshes. Let Ω_h denote the collection of mesh elements and $\mathcal{E}_h$ denote the collection of mesh edges. A polynomial degree $p_K \geq 1$ is assigned to each element and a degree p_E is assigned to each mesh edge E. For the first method, the practical trial space is $U_h = \Sigma_h \times W_h \times M_h \times F_h$ where

$$K_h = \{v : v|_K \in Q_{p_K, p_K}(K), \forall K \in \Omega_h\} \tag{7.1a}$$

$$\Sigma_h = (K_h)^3 \tag{7.1b}$$

$$W_h = (K_h)^2 \tag{7.1c}$$

$$M_h = \left\{\mu : \mu|_E \in P_{p_E+1}(E), \forall E \in \mathcal{E}_h \text{ and } \mu \in H_0^1(\Omega)\right\} \tag{7.1d}$$

$$F_h = \{\eta : \eta|_E \in P_{p_E}(E), \forall E \in \mathcal{E}_h\} \tag{7.1e}$$

(and, as before, $Q_{l,m}$ is the space of bivariate polynomials which are of degree at most l in x and m in y). Notice that in (7.1b), we interpret each element of $(K_h)^3$ as a symmetric matrix with entries in K_h (so the stress approximations are strongly symmetric). The edge order p_E is determined using the maximum rule, i.e. p_E is the maximum order of all elements adjacent to edge E.

Recall that the test space is determined by the trial-to-test operator $T : U \mapsto V$, which in turn requires inverting the Riesz map corresponding to the inner product in the test space. In practice, this is solved on the discrete level by using $\tilde{T} : U \mapsto \tilde{V}$ which is defined as

$$(\tilde{T}u, \tilde{v})_V = b(u, \tilde{v}) \qquad \forall\, \tilde{v} \in \tilde{V} \tag{7.2}$$

where $\tilde{V}$ is a finite dimensional subspace of V. For our implementation, $\tilde{V}$ is defined as

$$\tilde{V} = \left\{(\tau, v) : \tau|_K \in (Q_{\tilde{p}_K,\tilde{p}_K})^3 \text{ and } v|_K \in (Q_{\tilde{p}_K,\tilde{p}_K})^2\right\} \tag{7.3}$$

where $\tilde{p}_K = p_K + \delta p$. The default choice for the enrichment degree δp is 2. Numerical experience shows that this is a sufficient choice for most problems (see figure 5(d)).

Finally, we approximate the energy norm of the error using the error representation function $\tilde{e} \in \tilde{V}$ where $\tilde{e} = \tilde{T}(u - u_h)$. Note that by the definition of $\tilde{T}$, the error representation function can be computed element wise by solving the variational problem

$$(\tilde{e}, \tilde{v})_V = (\tilde{T}(u - u_h), \tilde{v})_V = b(u - u_h, \tilde{v}) = l(\tilde{v}) - b(u_h, \tilde{v}). \tag{7.4}$$

This implies that the energy norm of the error is approximated by

$$\|u - u_h\|_E = \sup_{v \in V} \frac{|b(u - u_h, v)|}{\|v\|_V} = \|T(u - u_h)\|_V \approx \|\tilde{T}(u - u_h)\|_V = \|\tilde{e}\|_V. \tag{7.5}$$

From Theorem 5.1, it is known that the energy error is equivalent to the standard norm error on U, so our choice of error indicator is justified assuming that the approximation of T by $\tilde{T}$ is sufficient.

In all cases, the standard test space norm is used for the inversion of the Riesz map, i.e.,

$$\|(\tau, v)\|_V^2 = \|\tau\|_{H(\mathrm{div},\Omega_h)}^2 + \|v\|_{H^1(\Omega_h)}^2. \tag{7.6}$$

Note that due to the strong symmetry of functions in the discrete space Σ_h, the term involving q in (2.5d) vanishes. Hence, at the discrete level, we may (omit all q's and) work with the reduced test space $V = H(\mathrm{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{V})$ instead of (2.5c). This is why we use the norm (7.6) in place of (2.8b).

7.2. **Test Problems.**

7.2.1. *Smooth Solution.* The first problem studied was a smooth solution with manufactured body force found by applying the elasticity equations to the exact displacements

$$u_x = \sin(\pi x)\sin(\pi y) \tag{7.7}$$

$$u_y = \sin(\pi x)\sin(\pi y) \tag{7.8}$$

over a unit square domain $\Omega = (0, 1) \times (0, 1)$ with u_x, u_y prescribed on $\partial\Omega$. For all cases, Ω_h is initially a uniform mesh of 4 square elements.

7.2.2. *L-Shaped Steel.* The second problem considered was the classical L-shaped domain problem with the material properties of steel. In polar coordinates r, θ around the re-entrant corner, the singular solution takes the form (see e.g., [33] or [24, § 4.2])

$$\sigma_r = r^{a-1}\left[F''(\theta) + (a+1)F(\theta)\right] \tag{7.9}$$

$$\sigma_\theta = a(a+1)r^{a-1}F(\theta) \tag{7.10}$$

$$\sigma_{r\theta} = -ar^{a-1}F'(\theta) \tag{7.11}$$

$$u_r = \frac{1}{2\mu}r^a\left[-(a+1)F(\theta) + \left(1-\frac{\nu}{1+\nu}\right)G'(\theta)\right] \tag{7.12}$$

$$u_\theta = \frac{1}{2\mu}r^a\left[-F'(\theta) + \left(1-\frac{\nu}{1+\nu}\right)(a-1)G(\theta)\right] \tag{7.13}$$

where $\nu = \frac{\lambda}{2(\lambda+\mu)}$ is Poisson's ratio and the functions $F(\theta)$ and $G(\theta)$ are given by

$$F(\theta) = C_1\sin(a+1)\theta + C_2\cos(a+1)\theta + C_3\sin(a-1)\theta + C_4\cos(a-1)\theta, \tag{7.14}$$

$$G(\theta) = \frac{4}{a-1}\left[-C_3\cos(a-1)\theta + C_4\sin(a-1)\theta\right]. \tag{7.15}$$

To determine the constants C_1, C_2, C_3, C_4 and a, we use the kinematic boundary conditions along the edges forming the reentrant corner (which without loss of generality we can take to be the edges $\theta = \pm 3\pi/4$). We can obtain a square integrable solution that satisfies $\operatorname{div}\sigma = 0$ by setting $C_2 = C_4 = 0$, $C_3 = 1$ and

$$C_1 = \frac{\left[4\left(1-\frac{\nu}{1+\nu}\right) - (a+1)\right]\sin\left((a-1)\frac{3}{4}\pi\right)}{(a+1)\sin\left((a+1)\frac{3}{4}\pi\right)} \tag{7.16}$$

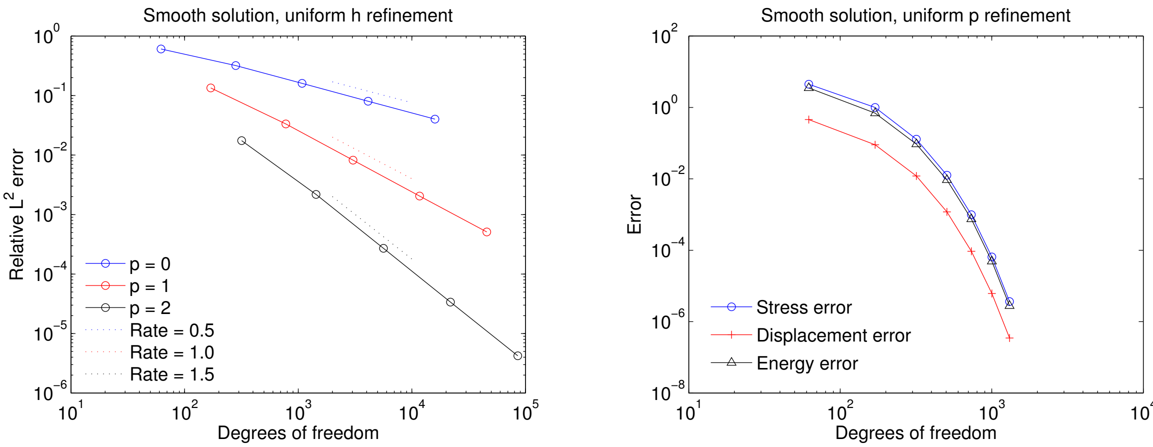
and letting $0 < a < 1$ be the solution of the transcendental equation

$$C_1\cos\left(\frac{3(a+1)\pi}{4}\right)(a+1) + \cos\left(\frac{3(a-1)\pi}{4}\right)(a-1) + 4\left(1-\frac{\nu}{1+\nu}\right)\cos\left(\frac{3(a-1)\pi}{4}\right) = 0.$$

Numerically solving for a with the material properties of steel [21], namely $\lambda = 123$ GPa, $\mu = 79.3$ GPa, we obtain $a \approx 0.6038$. This implies that all stress components have a singularity of strength (approximately) $r^{-0.3962}$ while the displacement components are smooth at the origin (but have singular derivatives). We can thus expect the stress components to be in (a space close to) $H^{0.6038-\epsilon}$ and the displacement components in $H^{1.6038-\epsilon}$ for $\epsilon > 0$.

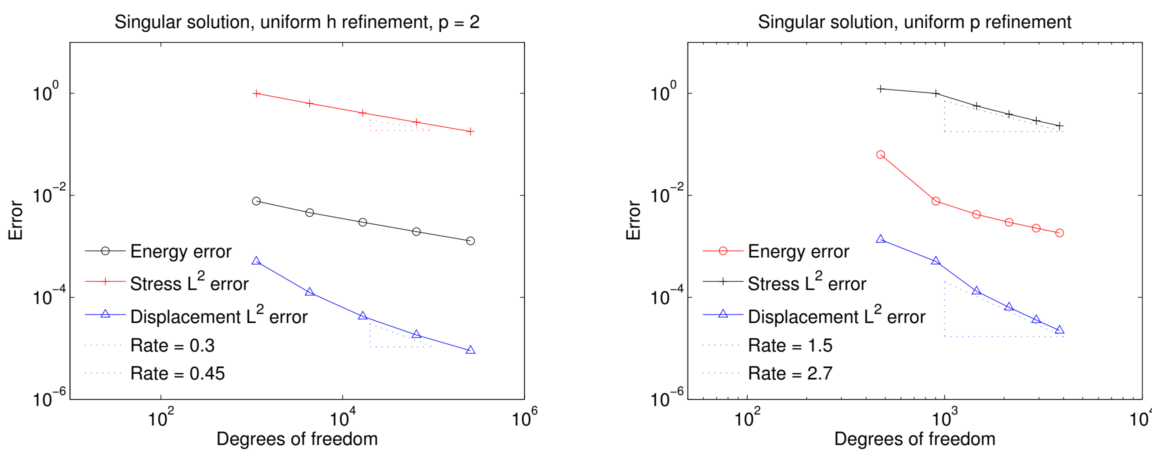
7.3. **Convergence rates.** The observed decrease of the error as the degrees of freedom increase is shown in Figure 1 for the smooth solution case and in Figure 2 for the L-shaped domain. Note that when we report the "L^2 error", we only consider the L^2-norm of the errors in u and σ, not the error in numerical fluxes or traces.

Consider the case of the smooth solution first. If we perform uniform h-refinements, the number of degrees of freedom (N) is $O(h^{-2})$. From Corollary 6.1, we expect to see the error decrease by $O(h^{p+1})$ for the smooth solution case, i.e., $O(N^{-(p+1)/2})$ in terms of N. This is confirmed in Figure 1(a). Also, since both displacement and stress are infinitely smooth, they converge at the same rate. For uniform p-refinements, exponential convergence is observed in Figure 1(b).



(a) Uniform h-refinements. (b) Uniform p-refinements.

FIGURE 1. Uniform refinement strategies for the smooth problem.



(a) Uniform h refinements. (b) Uniform p refinements.

FIGURE 2. Uniform refinement strategies for the L-shaped domain.

In the singular case of the L-shaped domain, Figure 2 shows the observed convergence history for uniform refinements. Since the stress variables are in $H^{0.6038-\epsilon}$, we expect that the best approximation error for stress should decrease at rate $h^{0.6038}$, or $N^{-0.3019}$. This is in agreement with Figure 2(a). Additionally, since displacement is in $H^{1.6038-\epsilon}$, one might think that its best approximation error should decrease more or less at rate $h^{1.6038}$, or $N^{-0.8019}$. However in the DPG method the errors for both these variables are coupled together. So, while we observe the optimal convergence rate for the stress variable, the convergence rate for the displacement seems to be somehow limited by the convergence rate of the stress. For uniform p-refinements, because we are considering a singular solution, the convergence rate is limited by the regularity of the solution, so unlike Figure 1(b), no exponential convergence is observed in Figure 2(b).

7.4. **Comparison with the weakly symmetric mixed method.** The smooth solution problem was also implemented using the weakly symmetric mixed element given in [5]. This

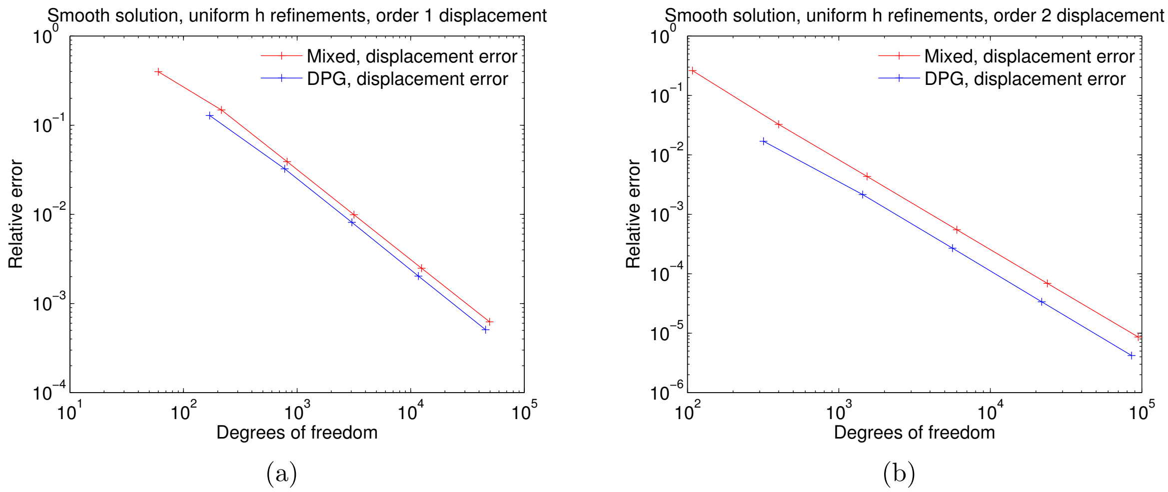


FIGURE 3. The DPG method vs. the mixed method.

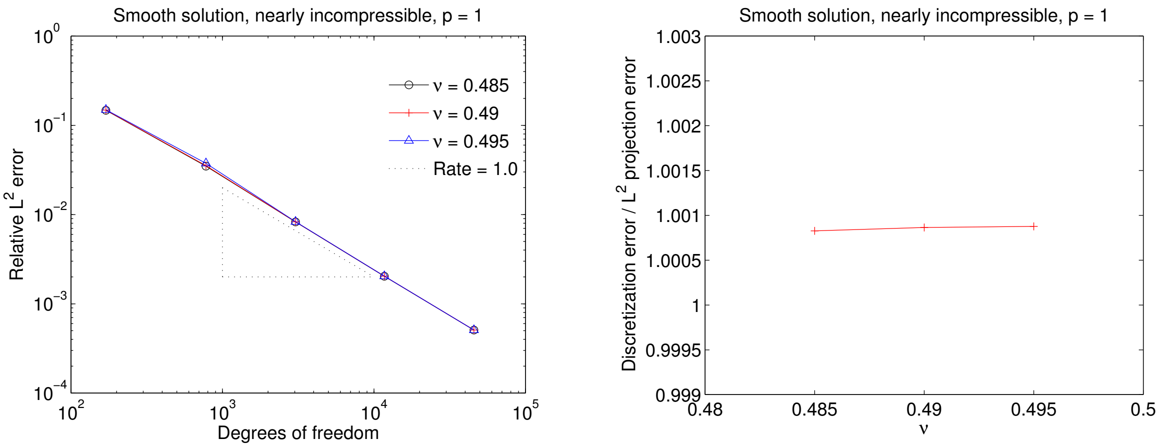
mixed method uses polynomials of degree one higher than our DPG method for the stress trial space. Expectedly therefore, the stress approximations given by the mixed method were generally observed to be superior in the L^2 norm. For the displacement however, both methods use the same space, so it is interesting to compare the displacement errors. This is done in Figure 3. The DPG method delivers lower displacement errors in the higher order case. In the lowest order case (not shown in the figure) the mixed method performs slightly better.

7.5. **Locking experiments.** In Figure 4 we show numerical evidence of the locking-free property of the DPG method. The figure shows convergence curves for various values of Poisson ratio close to the limiting value of 0.5. We used piecewise bilinear elements with homogeneous material data. The convergence curves in Figure 4(a) show hardly any difference as ν approaches 0.5. To be clearer, we also plot the ratio of the L^2 discretization error to the best approximation error (in σ and u combined) in Figure 4(b). We see that the ratio remains close to the optimal value of 1.0 even as ν approaches 0.5.

7.6. **Adaptivity.** All our adaptive schemes are based on the "greedy" strategy described in [13]. This means that all elements which contribute 50% of the maximum element error to the previously described error representation function are marked for refinement. For hp-adaptivity, we used the strategy suggested in [1], i.e., if an element contains the singularity, it is h-refined, otherwise it is p-refined.

Figure 5 shows results from both adaptivity schemes. Note that a nearly optimal rate of $O(N^{-1.2})$is observed for the h-adaptivity scheme. The hp-adaptive scheme results in an optimal rate of $O(N^{-1.5})$. Finally, Figure 5(e) shows the hp mesh obtained after 12 iterations and Figure 5(f) shows one component of the corresponding solution. The group relative L^2 error is reduced to 0.9%.

7.7. **Approximation of optimal test functions.** Figure 5(d) shows the effect of δp as seen in the h-adaptive process for the L-shaped domain problem. This measures the effect of approximating optimal test functions using the operator $\tilde{T}$ as opposed to T, Since the



(a) Convergence curves for various ν

(b) The ratio of the discretization error to the error in best approximation as a function of ν

FIGURE 4. Illustration of locking-free convergence (smooth data case).

curves for $\delta p = 2, 3, 4$ are coincident, it appears that we are sufficiently approximating the optimal test functions.

APPENDIX A. A PROPERTY OF THE WEAKLY SYMMETRIC MIXED FORMULATION

We consider a mixed method for linear elasticity with weakly imposed stress symmetry. The method we consider differs from a standard method [4] only in that it has an extra Lagrange multiplier. It is well known that the mixed formulation does not lock (see e.g.,[7, 8, 32]) for homogeneous isotropic material parameters. In this appendix, we will provide a stability result for slightly more general materials. Note however, that the main goal of this appendix is to establish stability estimates for the mixed method in the form needed for the analysis of the DPG scheme in the earlier sections.

The formulation reads as follows: Find $(\sigma, u, \rho, a) \in H(\mathrm{div}, \Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{K}) \times \mathbb{R}$ satisfying

$$(A\sigma, \tau)_\Omega + (u, \mathrm{div}\, \tau)_\Omega + (\rho, \tau)_\Omega + (aQ_0^{-1}AI, \tau)_\Omega = (F_1, \tau)_\Omega, \tag{A.1a}$$

$$(\mathrm{div}\, \sigma, v)_\Omega = (F_2, v)_\Omega, \tag{A.1b}$$

$$(\sigma, \eta)_\Omega = (F_3, q)_\Omega, \tag{A.1c}$$

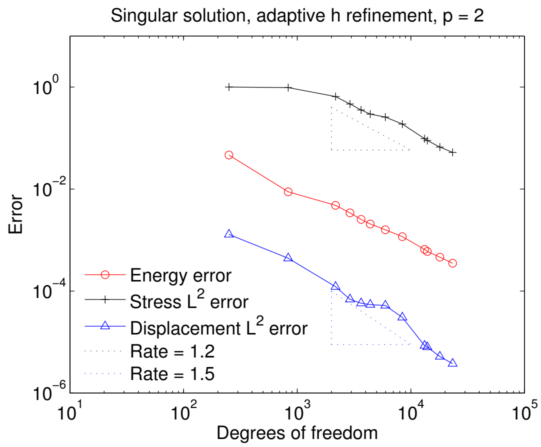
$$(\sigma, bQ_0^{-1}AI)_\Omega = (F_4, bI)_\Omega, \tag{A.1d}$$

for all $(\tau, v, \eta, b) \in H(\mathrm{div}, \Omega; \mathbb{M}), \times L^2(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{K}) \times \mathbb{R}$. Recall that Q_0 is as defined in (3.1). This formulation, specifically (A.1d), is motivated by the same constraint that motivated the second DPG method, namely (3.2).

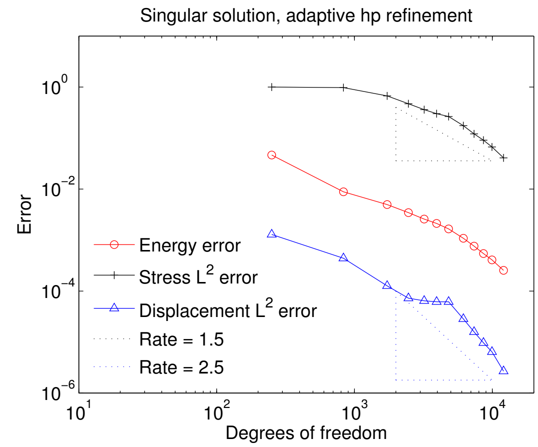
Theorem A.1. *Let* $(F_1, F_2, F_3, F_4) \in L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{M})$. *Then:*

(1) *Problem* (A.1) *is uniquely solvable and the solution component* u *is in fact in* $H_0^1(\Omega; \mathbb{V})$.
(2) *There is a positive constant* C_0 *such that*

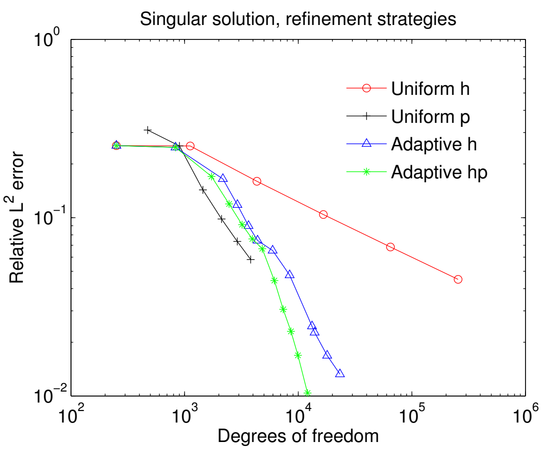
$$\|\sigma\|_{H(div,\Omega)} + \|u\|_{H^1(\Omega)} + \|\rho\|_\Omega + |a| \le C_0\big(\|F_1\|_\Omega + \|F_2\|_\Omega + \|F_3\|_\Omega + \|F_4\|_\Omega\big). \tag{A.2}$$



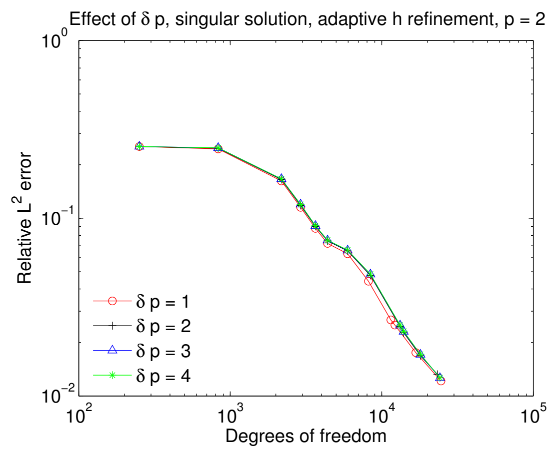
(a) Adaptive h refinements.



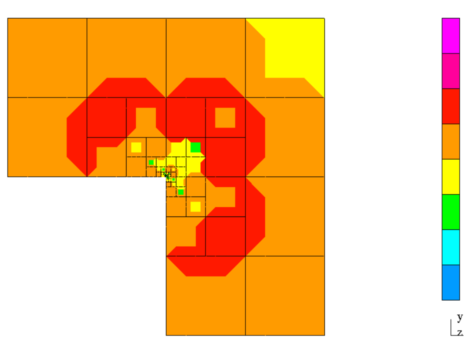
(b) Adaptive hp refinements.



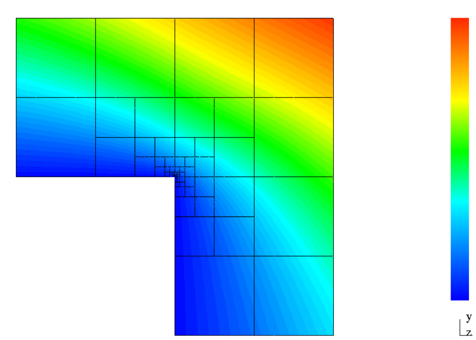
(c) A comparison of refinement strategies.



(d) Various optimal test function approximations.



(e) The hp mesh after 12 iterations. Element degrees are represented by color. (The color scale is tied to $p+1$.)



(f) The x-component of the computed displacement (u_x). The color scale indicates value of u_x.

FIGURE 5. Results from the adaptive scheme for the L-shaped domain.

(3) *In addition, if Assumption 3.1 holds, then the constant* C_0 *in* (A.2) *takes the form*

$$C_0 = \bar{c}_1 P_0^{-1} B^4 (\|A\| + P_0 + 1)^2 (\|A\| + B), \tag{A.3}$$

where $\bar{c}_1$ *is a positive constant independent of* A.

We will use the Babuška-Brezzi theory [7] and results from [4] to prove this theorem. In order to verify the conditions of the theory, we will use the following lemma.

Lemma A.2. *If Assumption 3.1 holds, then there is a positive constant* $\bar{c}_0$ *independent of the material coefficient* A *such that*

$$\|\operatorname{tr}\varphi\|_\Omega^2 \le \bar{c}_0^2 B^2 \left(\|\varphi_D\|_\Omega^2 + \|\operatorname{div}\varphi\|_\Omega^2\right), \tag{A.4}$$

for any $\varphi \in H(div, \Omega; \mathbb{M})$ *which satisfies*

$$\int_\Omega \operatorname{tr}(A\varphi) = 0. \tag{A.5}$$

Proof. This proof is similar to a proof in [7]. We can apply a standard regular right inverse of divergence to $\operatorname{tr}(A\varphi)$ since (A.5). Hence there exists a constant $c_0 > 0$ and $\eta \in H_0^1(\Omega; \mathbb{V})$ such that

$$\operatorname{div}\eta = Q_0^{-1}\operatorname{tr}(A\varphi), \quad \|\eta\|_{H^1(\Omega)} \le c_0 Q_0^{-1}\|\operatorname{tr}(A\varphi)\|_\Omega.$$

By the isotropy assumption – see (3.7) – we have that $\operatorname{tr}(A\varphi) = Q\operatorname{tr}\varphi$. This implies that

$$\operatorname{div}\eta = (QQ_0^{-1})\operatorname{tr}\varphi, \quad \|\eta\|_{H^1(\Omega)} \le c_0 Q_0^{-1}\|Q\operatorname{tr}\varphi\|_\Omega.$$

Then, since $QQ_0^{-1} \ge 1$ a.e., we have

$$\begin{aligned}
\|\operatorname{tr}\varphi\|_\Omega^2 &\le ((QQ_0^{-1})\operatorname{tr}\varphi, \operatorname{tr}\varphi)_\Omega = (\operatorname{div}\eta, \operatorname{tr}\varphi)_\Omega = ((\operatorname{div}\eta)I, \varphi)_\Omega \\
&= N(\nabla\eta - (\nabla\eta)_D, \varphi)_\Omega = -N(\eta, \operatorname{div}\varphi)_\Omega - N((\nabla\eta)_D, \varphi)_\Omega \\
&= -N(\eta, \operatorname{div}\varphi)_\Omega - N(\nabla\eta, \varphi_D)_\Omega \\
&\le N\|\eta\|_{H^1(\Omega)}\left(\|\varphi_D\|_\Omega^2 + \|\operatorname{div}\varphi\|_\Omega^2\right)^{1/2} \\
&\le c_0 N Q_0^{-1}\|Q\operatorname{tr}\varphi\|_\Omega\left(\|\varphi_D\|_\Omega^2 + \|\operatorname{div}\varphi\|_\Omega^2.\right)^{1/2}
\end{aligned}$$

Setting $\bar{c}_0 = c_0 N$, the lemma is proved. ☐

Proof of Theorem A.1. To apply the Babuška-Brezzi theory, we need to verify two conditions: (i) the coercivity on kernel, and (ii) the inf-sup condition.

Step (i). *Coercivity on kernel:* Define the kernel space

$$V_0 = \{\tau \in H(\operatorname{div}, \Omega; \mathbb{M}) : \operatorname{div}\tau = 0,\ \tau' = \tau,\ \int_\Omega \operatorname{tr}(A\tau) = 0\}.$$

Clearly, if A is uniformly coercive, then there is a positive constant c_1, depending on A, such that

$$c_1\|\tau\|_{H(\operatorname{div},\Omega)}^2 \le (A\tau, \tau)_\Omega, \qquad \forall\tau \in V_0. \tag{A.6}$$

If in addition Assumption 3.1 holds, then we can give the dependence of c_1 on P_0 and Q, as we see now. For any $\tau \in V_0$,

$$(A\tau, \tau)_\Omega \ge (P\tau_D, \tau_D)_\Omega \ge P_0\|\tau_D\|_\Omega^2.$$

Using Lemma A.2 and the fact that $\operatorname{div}\tau = 0$, we have that

$$\|\tau_D\|_\Omega^2 \ge (\bar{c}_0 B)^{-2}\|\operatorname{tr}\tau\|_\Omega^2.$$

Since $\|\tau\|_\Omega^2 = \|\tau_D\|_\Omega^2 + N^{-1}\|\operatorname{tr}\tau\|_\Omega^2$, we have that

$$\|\tau\|_\Omega^2 \le (1 + N^{-1}\bar{c}_0^2 B^2)\|\tau_D\|_\Omega^2.$$

So, for any $\tau \in V_0$, we have that

$$\frac{P_0}{1+N^{-1}\bar{c}_0^2 B^2}\|\tau\|_{H(\operatorname{div},\Omega)}^2 = \frac{P_0}{1+N^{-1}\bar{c}_0^2 B^2}\|\tau\|_\Omega^2 \le (A\tau,\tau)_\Omega$$

and we conclude that (A.6) holds with

$$c_1 = \frac{P_0}{1+\bar{c}_0^2 B^2} \tag{A.7}$$

in the isotropic case.

Step (ii). *Inf-sup condition:* The inf-sup condition will follow once we show that there is a positive constant c_2 such that for any $(u,\rho,a) \in L^2(\Omega;\mathbb{V}) \times L^2(\Omega;\mathbb{K}) \times \mathbb{R}$, there is a $\tau \in H(\operatorname{div},\Omega;\mathbb{M})$ satisfying

$$(u,\operatorname{div}\tau)_\Omega + (\rho,\tau)_\Omega + (aQ_0^{-1}AI,\tau)_\Omega \ge c_2\|\tau\|_{H(\operatorname{div},\Omega)}\left(\|u\|_\Omega + \|\rho\|_\Omega + |a|\right). \tag{A.8}$$

To this end, we first recall [3, Theorem 11.1]. Accordingly, there is a $\tau_0 \in H(\operatorname{div},\Omega;\mathbb{M})$ and $c_3 > 0$ such that $\operatorname{div}\tau_0 = u$, $\operatorname{skw}\tau_0 = \rho$, and

$$\|\tau_0\|_{H(\operatorname{div},\Omega)} \le c_3(\|u\|_\Omega + \|\rho\|_\Omega). \tag{A.9}$$

The constant c_3 depends only on Ω.

To prove (A.8), we choose τ of the form $\tau = \tau_0 + \lambda I$ where $\lambda \in \mathbb{R}$. Obviously,

$$\tau \in H(\operatorname{div},\Omega,\mathbb{M}), \quad \operatorname{div}\tau = u, \quad \text{and} \quad \operatorname{skw}\tau = \rho,$$

for any $\lambda \in \mathbb{R}$. So, to show the estimate (A.8), we need only choose $\lambda \in \mathbb{R}$ such that

$$(aQ_0^{-1}AI,\tau)_\Omega = (aQ_0^{-1}AI,\tau_0+\lambda I)_\Omega = |a|^2,$$

i.e.,

$$\lambda = \frac{a - Q_0^{-1}(AI,\tau_0)_\Omega}{Q_0^{-1}\int_\Omega \operatorname{tr}(AI)}. \tag{A.10}$$

Then, by (A.9), there is a positive constant c_2 such that (A.8) holds.

If the material is isotropic, then the dependence of c_2 on the components of A can be tracked, as follows. Observe that since $QQ_0^{-1} \ge 1$, we have

$$Q_0^{-1}\int_\Omega \operatorname{tr}(AI) = N\int_\Omega Q_0^{-1}Q \ge N|\Omega|,$$

$$|Q_0^{-1}(AI,\tau_0)_\Omega| = \left|\int_\Omega (Q_0^{-1}Q)\operatorname{tr}\tau_0\right| \le c_4 B\|\tau_0\|_\Omega,$$

with a constant c_4 depending only on Ω. Using this in (A.10), we have

$$\begin{aligned}\|\tau\|_{H(\operatorname{div},\Omega)} &\le \|\tau_0\|_{H(\operatorname{div},\Omega)} + \|\lambda I\|_{H(\operatorname{div},\Omega)}\\ &\le c_3(\|u\|_\Omega + \|\rho\|_\Omega) + \frac{c_3c_4B(\|u\|_\Omega+\|\rho\|_\Omega)+|a|}{\sqrt{N|\Omega|}}.\end{aligned}$$

Thus there is a constant c_5 depending only on Ω such that

$$\|\tau\|_{H(\operatorname{div},\Omega)} \le c_5 B(\|u\|_\Omega + \|\rho\|_\Omega + |a|),$$

and therefore (A.8) holds with

$$c_2 = (c_5 B)^{-1} \tag{A.11}$$

in the isotropic case.

Step (iii). Since we have verified the two conditions of the Babuška-Brezzi theory, we conclude that there is a unique solution $(\sigma, u, \rho, a) \in H(\mathrm{div}, \Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{V}) \times L^2(\Omega; \mathbb{K}) \times \mathbb{R}$. Moreover, the theory guarantees – see [7, Eq. (1.29)–(1.30)] – that the stability estimate

$$\|\sigma\|_{H(\mathrm{div},\Omega)} + \|u\|_\Omega + \|\rho\| + |a| \le \tilde{C}(\|F_1\|_\Omega + \|F_2\|_\Omega + \|F_3\|_\Omega + \|F_4\|_\Omega)$$

holds with

$$\tilde{C} = c_1^{-1} c_2^{-2} (\|A\| + c_1 + c_2)^2.$$

If in addition, Assumption 3.1 holds, then by (A.7) and (A.11),

$$\tilde{C} = P_0^{-1}(1 + \bar{c}_0^2 B^2) c_5^2 B^2 (\|A\| + (c_5 B)^{-1} + P_0)^2. \tag{A.12}$$

Step (iv). To prove that u is in fact in $H^1(\Omega, \mathbb{V})$, we observe that by choosing $\tau \in \mathcal{D}(\Omega; \mathbb{M})$ arbitrarily, we can conclude that the equality

$$A\sigma - \nabla u + \rho + a Q_0^{-1} A I = F_1 \tag{A.13}$$

holds in the sense of distributions. Hence $u \in H^1(\Omega; \mathbb{V})$. Consequently, we may integrate (A.1a) by parts for any $\tau \in H(\mathrm{div}, \Omega; \mathbb{M})$ and use (A.13) to conclude that $u \in H_0^1(\Omega; \mathbb{V})$. By (A.13),

$$\|\sigma\|_{H(\mathrm{div},\Omega)} + \|u\|_{H^1(\Omega)} + \|\rho\| + |a| \le \tilde{C}(2 + \|A\| + Q_0^{-1}\|AI\|_\Omega)(\|F_1\|_\Omega + \|F_2\|_\Omega + \|F_3\|_\Omega + \|F_4\|_\Omega).$$

We have thus proved (A.2) with $C_0 = \tilde{C}(2 + \|A\| + Q_0^{-1}\|AI\|_\Omega)$.

If in addition, the material is isotropic, then by (A.12), the constant C_0 can be written as

$$C_0 = P_0^{-1}(1 + \bar{c}_0^2 B^2) c_5^2 B^2 (\|A\| + (c_5 B)^{-1} + P_0)^2 (2 + \|A\| + N|\Omega| B).$$

Since $B \ge 1$, we conclude that there is a positive constant $\bar{c}_1$ such that $C_0 \le \bar{c}_1 P_0^{-1} B^4 (\|A\| + P_0 + 1)^2 (\|A\| + B)$, as stated in (A.3). ☐

References

[1] M. Ainsworth and B. Senior, *Aspects of an adaptive hp-finite element method: adaptive strategy, conforming approximation and efficient solvers*, Comput. Methods Appl. Mech. Engrg., 150 (1997), pp. 65–87. Symposium on Advances in Computational Mechanics, Vol. 2 (Austin, TX, 1997).

[2] D. N. Arnold, G. Awanou, and R. Winther, *Finite elements for symmetric tensors in three dimensions*, Math. Comp., 77 (2008), pp. 1229–1251.

[3] D. N. Arnold, R. S. Falk, and R. Winther, *Finite element exterior calculus, homological techniques, and applications*, Acta Numer., (2006), pp. 1–155.

[4] ——, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Math. Comp., 76 (2007), pp. 1699–1723 (electronic).

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

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

[7] F. Brezzi, and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, ISBN-10: 3540975829 (1991).

[8] C. Carstensen, G. Dolzmann, S. A. Funken, and D. S. Helm, *Locking-free adaptive mixed finite element methods in linear elasticity*, Comput. Methods Appl. Mech. Engrg., 190 (2000), pp. 1701–1718.

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

[10] L. Demkowicz. *Computing with hp Finite Elements. I. One- and Two-Dimensional Elliptic and Maxwell Problems.* Chapman & Hall/CRC Press, Taylor and Francis, October 2006.

[11] L. Demkowicz, *Polynomial exact sequences and projection-based interpolation with application to maxwell equations*, in Lecture Notes in Mathematics, Springer-Verlag, 2008.

[12] L. Demkowicz and A. Buffa, *H^1, $H(\mathrm{curl})$ and $H(\mathrm{div})$-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 267–296.

[13] L. Demkowicz and J. Gopalakrishnan, Analysis of the DPG method for the Poisson equation. *SIAM J Numer. Anal.*, 49(5):1788–1809, 2011.

[14] ——, *A class of discontinuous Petrov-Galerkin methods. Part I: The transport equation*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1558–1572.

[15] ——, *A class of discontinuous Petrov-Galerkin methods. Part II: Optimal test functions*, Numerical Methods for Partial Differential Equations, 27 (2011), pp. 70–105.

[16] L. Demkowicz, J. Gopalakrishnan, and A. Niemi, A class of discontinuous Petrov-Galerkin methods. Part III: Adaptivity. *Applied Numerical Mathematics*, 62:396–427, 2012.

[17] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl, *Polynomial extension operators. Part I*, SIAM J. Numer. Anal., 46 (2008), pp. 3006–3031.

[18] ——, *Polynomial extension operators. Part II*, SIAM J. Numer. Anal., 47 (2009), pp. 3293–3324.

[19] ——, Polynomial extension operators. Part III. *Math. Comp.*, 81(279):1289–1326, 2012.

[20] R. S. Falk, *Finite element methods for linear elasticity*, in Mixed Finite Elements, Compatibility Conditions, and Applications, pp. 160–194. Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy, June 26-July 1, 2006,.

[21] Y. C. Fung and P. Tong. Classical and Computational Solid Mechanics. World Scientific Publishing Co., 2001.

[22] J. Gopalakrishnan and J. Guzmán, A second elasticity element using the matrix bubble. *IMA J. Numer. Anal.*, 32:352–372, 2012.

[23] J. Gopalakrishnan and J. Guzmán, Symmetric non-conforming mixed finite elements for linear elasticity. *SIAM J Numer. Anal.*, 49(4):1504–1520, 2011.

[24] P. Grisvard, *Singularities in boundary value problems*, vol. 22 of Recherches en Mathématiques Appliquées [Research in Applied Mathematics], Masson, Paris, 1992.

[25] W. W. Hager, *Updating the inverse of a matrix*, SIAM Rev., 31 (1989), pp. 221–239.

[26] C. O. Horgan, *Korns inequalities and their applications in continuum mechanics*, SIAM Review, 37 (1995), pp. 491–511.

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

[28] J. E. Marsden and T. J. Hughes, *Mathematical foundations of elasticity*, Dover, 1993.

[29] W. Qiu and L. Demkowicz, *Mixed hp-finite element method for linear elasticity with weakly imposed symmetry*, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 3682–3701.

[30] W. Qiu and L. Demkowicz, *Mixed hp-finite element method for linear elasticity with weakly imposed symmetry: stability analysis*, SIAM J. Numer. Anal., 49 (2011), pp. 619–641.

[31] 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.

[32] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math., 53 (1988), pp. 513–538.

[33] D. Vasilopoulos, *On the determination of higher order terms of singular elastic stress fields near corners*, Numer. Math., 53 (1988), pp. 51–95.

[34] J. Zitelli, I. Muga, L. Demkowicz, J. Gopalakrishnan, D. Pardo, and V. Calo, *A class of discontinuous Petrov-Galerkin methods. Part IV: Wave propagation*, Journal of Computational Physics, 230 (2011), pp. 2406–2432.

The Institute for Computational Engineering and Sciences, the University of Texas at Austin, Austin, Texas 78712
E-mail address: jbramwell@ices.utexas.edu

The Institute for Computational Engineering and Sciences, the University of Texas at Austin, Austin, Texas 78712
E-mail address: leszek@ices.utexas.edu

Department of Mathematics and Statistics, Portland State University, Portland OR 97207-0751
E-mail address: gjay@pdx.edu

The Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, Minnesota 55455
E-mail address: qiuxa001@ima.umn.edu