SIAM J. NUMER. ANAL.
Vol. 52, No. 3, pp. 1335–1353

A POSTERIORI ERROR CONTROL FOR DPG METHODS*

CARSTEN CARSTENSEN†, LESZEK DEMKOWICZ‡, AND JAY GOPALAKRISHNAN§

Abstract. A combination of ideas in least-squares finite element methods with those of hybridized methods recently led to discontinuous Petrov–Galerkin (DPG) finite element methods. They minimize a residual inherited from a piecewise ultraweak formulation in a nonstandard, locally computable, dual norm. This paper establishes a general a posteriori error analysis for the natural norms of the DPG schemes under conditions equivalent to a priori stability estimates. It is proven that the locally computable residual norm of *any discrete function* is a lower and an upper error bound up to explicit data approximation errors. The presented abstract framework for a posteriori error analysis applies to known DPG discretizations of Laplace and Lamé equations and to a novel DPG method for the stress-velocity formulation of Stokes flow with symmetric stress approximations. Since the error control does not rely on the discrete equations, it applies to inexactly computed or otherwise perturbed solutions within the discrete spaces of the functional framework. Numerical illustrations show that the error control is practically feasible.

Key words. least-squares, inexact solution, discontinuous, Petrov–Galerkin, DPG, Laplacian, linear elasticity, Stokes, true fluid stress

AMS subject classification. 65N30

DOI. 10.1137/130924913

1. Introduction.

The discontinuous Petrov–Galerkin (DPG) method of [15, 16] can simultaneously be viewed either as a finite element method with nonstandard test spaces or as a least-squares finite element method (LSFEM) minimizing the residual from an ultraweak formulation in a nonstandard norm. This paper, adopting the LSFEM point of view, introduces an a posteriori error analysis for the DPG method and establishes the efficiency and reliability of an a posteriori error estimator for the DPG method. This partly justifies the adaptive computations in [19] and identifies a data approximation error. Under the same conditions which guarantee stability and a priori error estimates, a novel abstract analysis is presented, which applies to the known DPG method for Laplace and Lamé equations as well to a novel DPG formulation for the Stokes equations. The a posteriori error estimator does *not* rely on Galerkin orthogonality in the sense that, as in LSFEM, the reliability holds for *all* approximations and, in particular, allows for *inexact solve*.

The functional framework for the method consists of a reflexive real Banach space $(X, \|\cdot\|_X)$ and a real Hilbert space $(Y, \langle\cdot,\cdot\rangle_Y)$ with associated norm $\|\cdot\|_Y = \langle\cdot,\cdot\rangle_Y^{1/2}$ and dual Y^* plus a continuous invertible linear operator $B : X \to Y^*$. Given any $F \in Y^*$, there exists a unique solution $x := B^{-1}F$ of

$$Bx = F. \tag{1.1}$$

Given any finite-dimensional subspace $X_h \subset X$, the *idealized DPG* method computes

*Received by the editors June 13, 2013; accepted for publication (in revised form) April 1, 2014; published electronically June 5, 2014. This work was partially supported by the NSF under grant DMS-1318916 and by the AFOSR under grant FA9550-12-1-0484.

http://www.siam.org/journals/sinum/52-3/92491.html

†Department of Mathematics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany (cc@math.hu-berlin.de). The first author thankfully acknowledges the support of a JTO Fellowship and the kind hospitality of the ICES at the University of Texas at Austin.

‡Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712 (leszek@ices.utexas.edu).

§Portland State University, Portland, OR 97207-0751 (gjay@pdx.edu).

an approximation

$$\widehat{x}_h = \arg\min_{\xi_h \in X_h} \|F - B\xi_h\|_{Y^*} \tag{1.2}$$

to x by minimizing the residual $\|F - B\xi_h\|_{Y^*}$ in X_h. Although the DPG method was originally presented through the concept of optimal test functions, the least-squares characterization (1.2) of the method has been noted in [16]. Many well-known least-squares Galerkin methods [5] have Y set to L^2. In such cases, the evaluation of the Y^*-norm of the residual in (1.2) is standard and reduces to computation of an integral.

An essential difference in the DPG methods considered in the applications below is the use of a space Y different from L^2. As shown in the earlier papers [7, 17, 16], it is possible to use hybridization ideas to recast a boundary value problem into an ultraweak variational formulation, employing a test space Y of functions with no continuity constraints across mesh element interfaces. This localizes the evaluation of the Y^*-norm in (1.2) to decoupled calculations on mesh elements. Even if so localized, the subspace of functions in Y supported on a single element is infinite dimensional. This motivates the *practical DPG method* with finite-dimensional subspace $Y_h \subset Y$: compute an approximation

$$x_h = \arg\min_{\xi_h \in X_h} \|F - B\xi_h\|_{Y_h^*} \tag{1.3}$$

to x by minimizing the residual $\|F - B\xi_h\|_{Y_h^*}$ in X_h.

The minimization problems (1.2)–(1.3) can be recast in terms of the bilinear form

$$b(x, y) := (Bx)(y) \qquad \text{for all } x \in X \text{ and } y \in Y \tag{1.4}$$

with the duality between Y^* and Y in the right-hand side. Then (1.1) reads $b(x, y) = F(y)$ for all $y \in Y$. The *ideal DPG* method (1.2) characterizes $\widehat{x}_h$ in X_h as the solution to

$$b(\widehat{x}_h, y) = F(y) \qquad \text{for all } y \in \widehat{Y}_h^{\text{opt}} := T(X_h),$$

where $T : X \to Y$ is defined by $\langle Tx, y\rangle_Y = b(x, y)$ for all $x \in X$ and $y \in Y$. Similarly, the *practical DPG* method (1.3) solves

$$b(x_h, y) = F(y) \qquad \text{for all } y \in Y_h^{\text{opt}} := T_h(X_h) \tag{1.5}$$

for $x_h \in X_h$, where $T_h : X \to Y_h$ defined by $\langle T_h x, y\rangle_Y = b(x, y)$ for all $x \in X$ and $y \in Y_h$. The idea to reformulate (1.5) as a mixed formulation involving x_h and the corresponding residual is explored in [13]. The Petrov–Galerkin character of the ideal and practical DPG schemes is clearly evident from these reformulations.

The practical computation of the DPG solution x_h in (1.3) reduces to the solution of a symmetric positive definite linear system. One interpretation of the difference between the ideal and the practical DPG methods is that the test space is *inexactly* computed in the practical version (using T_h instead of the ideal T). The calculation of Tx requires an exact solve of a problem in Y, while the calculation of $T_h x$ solely solves some problem in the finite-dimensional subspace $Y_h \subset Y$.

This paper establishes an a posteriori error analysis of the practical DPG method (1.3). The a posteriori error analysis of conforming finite element schemes is well developed [2, 3, 4, 20, 32] and even nonstandard finite element technologies are understood [9, 11]. The residual norm

$$\eta = \|F - Bx_h\|_{Y_h^*} \tag{1.6}$$

equals the norm of the *approximate error representation function* $\varepsilon_h \in Y_h$ of [19] computed from

$$\langle \varepsilon_h, y\rangle_Y = F(y) - b(x_h, y) \qquad \text{for all } y \in Y_h. \tag{1.7}$$

The Riesz representation theorem implies

$$\eta = \|\varepsilon_h\|_Y. \tag{1.8}$$

Under certain conditions of [22] on Y_h and Y, this paper establishes reliable and efficient a posteriori error control

$$C_2\,\eta \leq \|x - x_h\|_X \leq C_3\,\eta + C_4 \operatorname{osc}(F) \tag{1.9}$$

with mesh-independent constants C_2, C_3, C_4 and the data approximation term $\operatorname{osc}(F)$ defined below in (1.12). In the applications, this data approximation term will be a data oscillation and of higher order for piecewise smooth data. The estimates of (1.9) follow from Theorem 2.1 below. In fact, the theorem provides such an estimate not only for the minimizer x_h, but more generally for *any* $\tilde{x}_h \in X_h$ and corresponding $\tilde{\eta}$. Similar properties are shared by LSFEM.

One of the conditions in the a priori convergence analysis of [22] is the existence of a bounded linear operator $\Pi : Y \to Y_h$ with the property

$$b(X_h, (I - \Pi)Y) = 0. \tag{1.10}$$

In other words, for any $\xi_h \in X_h$ and for any $y \in Y$, $b(\xi_h, y - \Pi y) = 0$. It is shown in [22] that this Fortin-type operator Π guarantees discrete stability and quasi-optimal convergence of the practical DPG method in the sense that

$$\|x - x_h\|_X \leq C_1 \min_{\xi_h \in X_h} \|x - \xi_h\|_X \tag{1.11}$$

holds with a mesh-independent constant C_1. This paper proves that (1.9) holds if the data approximation term is defined using Π by

$$\operatorname{osc}(F) := \|F \circ (1 - \Pi)\|_{Y^*}, \tag{1.12}$$

even when x_h is perturbed or inexactly computed.

The remaining parts of this paper are organized as follows. A set of abstract sufficient conditions in section 2 led to the equivalence (1.9) between the error and its estimator. The data approximation term is always efficient and results in an alternative proof of the a priori error estimate (1.11). Section 3 elucidates the a posteriori error analysis of the ultraweak formulation for the Laplace and Lamé equation and their DPG discretizations. The same section presents a new DPG method for Stokes flow featuring symmetric true fluid stress approximations. Adaptivity for these three model problems using discontinuous spaces has been heavily studied in the literature [1, 6, 24, 25, 26, 27, 28, 31] (see the review [10] for a more exhaustive bibliography). Section 4 provides numerical evidence for reliability and efficiency and shows that the estimator is useful in practice. Finally, as a notational foreword, standard notations from functional analysis, and for Lebesgue and Sobolev spaces, applies throughout this paper.

2. The abstract error control. This self-contained section is devoted to the abstract framework of DPG methods and its a posteriori error analysis based on three ingredients (H1)–(H3).

(H1) Suppose $b : X \times Y \to \mathbb{R}$ is a bounded real bilinear form in the reflexive real Banach space $(X, \|\cdot\|_X)$ and the real Hilbert space $(Y, \langle\cdot,\cdot\rangle_Y)$ and set

$$\|b\| := \sup_{x\in X\setminus\{0\}} \sup_{y\in Y\setminus\{0\}} \frac{b(x,y)}{\|x\|_X\, \|y\|_Y} < \infty.$$

(H2) Suppose b satisfies the inf-sup condition

$$0 < \beta := \inf_{x\in X\setminus\{0\}} \sup_{y\in Y\setminus\{0\}} \frac{b(x,y)}{\|x\|_X\, \|y\|_Y}$$

and the uniqueness condition $\{y \in Y : b(x,y) = 0\; \forall x \in X\} = \{0\}$.

(H3) The closed subspaces $X_h \subset X$ and $Y_h \subset Y$ admit a bounded linear operator $\Pi : Y \to Y_h$ with operator norm $\|\Pi\|$ and the property

$$b(X_h, (1-\Pi)Y) = 0.$$

(This notation abbreviates $b(\xi_h, (1-\Pi)y) = 0$ for all $\xi_h \in X_h$ and for all $y \in Y$.)

Let x_h and η denote the DPG solution and its error estimator η, given by (1.3) and (1.6), respectively. The main result below is stated for an *inexact solution* $\tilde{x}_h$, which is *any* discrete ansatz function in X_h. It shows that the corresponding error estimator $\tilde{\eta}$ provides *reliable* (i.e., $\tilde{\eta}$ provides an upper bound for the error) and *efficient* (i.e., $\tilde{\eta}$ provides a lower bound for the error) error control.

THEOREM 2.1 (reliability and efficiency). *Assume* (H1)–(H3). *Let* $F \in Y^*$, $x = B^{-1}F$, *and* $\tilde{x}_h \in X_h$. *Then the residual* $\tilde{\eta} := \|F - B\tilde{x}_h\|_{Y_h^*}$ *and the data-approximation error* $\operatorname{osc}(F) := \|F \circ (1-\Pi)\|_{Y^*}$ *satisfy*

$$\beta^2 \|x - \tilde{x}_h\|_X^2 \le \tilde{\eta}^2 + \Big(\|\Pi\|\, \tilde{\eta} + \operatorname{osc}(F)\Big)^2, \tag{2.1}$$

$$\tilde{\eta} \le \|b\|\, \|x - \tilde{x}_h\|_X, \quad \textit{and} \tag{2.2}$$

$$\operatorname{osc}(F) \le \|b\|\, \|1-\Pi\| \min_{\xi_h\in X_h} \|x - \xi_h\|_X. \tag{2.3}$$

Several comments are in order before the proof of the theorem concludes this section.

Remark 2.2 (bijectivity of B). It is well known that the operator $B : X \to Y^*$ in (1.4) is bijective under the conditions (H1)–(H2) and $\|B^{-1}\| = 1/\beta$.

Remark 2.3 (constants in error bounds). The a posteriori upper and lower error bounds (1.9) follow from Theorem 2.1 with $C_2 = 1/\|b\|$, $C_3 = (1 + \|\Pi\|^2)^{1/2}/\beta$, and $C_4 = 1/\beta$.

Remark 2.4 (data approximation error). The data approximation term

$$\|F \circ (1-\Pi)\|_{Y^*}$$

will be of higher order in all the applications of this paper. However, it is always strictly efficient in the sense of (2.3).

Remark 2.5 (alternative proof of quasi-optimal error estimates). Theorem 2.1 provides an alternative proof of the best approximation property (1.11). In fact, (1.3) implies, for any $\xi_h \in X_h$, that

$$\eta = \|F - Bx_h\|_{Y_h^*} \le \|B(x - \xi_h)\|_{Y_h^*} \le \|b\|\, \|x - \xi_h\|_X.$$

The combination of this with (2.1) and (2.3) implies

$$\beta\|x - x_h\|_X \le \|b\|\,(C_3 + C_4\,\|b\|\,\|1 - \Pi\|)\min_{\xi_h \in X_h}\|x - \xi_h\|_X.$$

This is (1.11) with $C_1 = C_3 + C_4\,\|b\|\,\|1 - \Pi\|$. The direct a priori error analysis of [22, Theorem 2.1] yields the sharper constant $C_1 = \|\Pi\|\,\|b\|/\beta$.

Remark 2.6 (stability is equivalent to (H3)). The aforementioned proof of a priori convergence indicates the role of (H3) as a sufficient condition for the convergence and hence for the stability of the DPG scheme. It is easy to see that the converse also holds: (H3) is necessary for stability of the DPG method.

Proof of Theorem 2.1. The proof of the reliability estimate (2.1) involves the error $e := x - \tilde{x}_h$, the error representation function $\tilde{\varepsilon}$ in Y defined by

$$\langle\tilde{\varepsilon}, y\rangle_Y = F(y) - b(\tilde{x}_h, y) \quad \text{for all } y \in Y, \tag{2.4}$$

and its approximation $\tilde{\varepsilon}_h$ in Y_h defined by

$$\langle\tilde{\varepsilon}_h, y_h\rangle_Y = F(y_h) - b(\tilde{x}_h, y_h) \quad \text{for all } y \in Y_h. \tag{2.5}$$

The reliability of $\tilde{\varepsilon}$ immediately follows from the inf-sup condition (H2),

$$\beta\|e\|_X \le \|Be\|_{Y^*} = \|\tilde{\varepsilon}\|_Y. \tag{2.6}$$

To prove the reliability of $\tilde{\varepsilon}_h$, observe that (2.4)–(2.5) imply

$$\delta := \tilde{\varepsilon} - \tilde{\varepsilon}_h \perp Y_h,$$

where $\perp$ denotes orthogonality with respect to the scalar product in Y. The Pythagorean theorem gives

$$\|\tilde{\varepsilon}\|_Y^2 = \|\tilde{\varepsilon}_h\|_Y^2 + \|\delta\|_Y^2. \tag{2.7}$$

Using Π from (H3), with the observation that $\Pi\delta \in Y_h \perp \delta$,

$$\begin{aligned}\|\delta\|_Y^2 &= \langle\delta, \delta - \Pi\delta\rangle_Y = \langle\tilde{\varepsilon} - \tilde{\varepsilon}_h, \delta - \Pi\delta\rangle_Y\\ &= \langle\tilde{\varepsilon}, \delta - \Pi\delta\rangle_Y + \langle\tilde{\varepsilon}_h, \Pi\delta\rangle_Y.\end{aligned}$$

Since $b(\tilde{x}_h, \delta - \Pi\delta) = 0$ by (H3),

$$\langle\tilde{\varepsilon}, \delta - \Pi\delta\rangle_Y = b(x - \tilde{x}_h, \delta - \Pi\delta) = F(\delta - \Pi\delta).$$

Combining with the previous identity,

$$\begin{aligned}\|\delta\|_Y^2 &= F(\delta - \Pi\delta) + \langle\tilde{\varepsilon}_h, \Pi\delta\rangle_Y\\ &\le (\|F \circ (1 - \Pi)\|_{Y^*} + \|\tilde{\varepsilon}_h\|_Y\,\|\Pi\|)\,\|\delta\|_Y.\end{aligned}$$

In other words,

$$\|\delta\|_Y \le \|F \circ (1 - \Pi)\|_{Y^*} + \|\tilde{\varepsilon}_h\|_Y\,\|\Pi\|. \tag{2.8}$$

The inequalities (2.6) and (2.8), together with the identity (2.7), prove

$$\beta^2\|e\|_X^2 \le \|\tilde{\varepsilon}_h\|_Y^2 + (\|F \circ (1 - \Pi)\|_{Y^*} + \|\tilde{\varepsilon}_h\|_Y\,\|\Pi\|)^2.$$

Since $\|\tilde{\varepsilon}_h\|_Y = \|F - B\tilde{x}_h\|_{Y_h^*} = \tilde{\eta}$ this proves the reliability estimate (2.1).

The efficiency estimate (2.2) immediately follows from (H1),

$$\tilde{\eta} = \|Be\|_{Y_h^*} \leq \|b\| \, \|e\|_X.$$

It only remains to prove (2.3). Considering any $y \in Y$ with $\|y\|_Y = 1$ and employing (H3),

$$\begin{aligned}(F \circ (1 - \Pi))\,(y) &= F(y - \Pi y) = b(x, y - \Pi y) \\ &= b(x - \xi_h, y - \Pi y) \leq \|b\| \, \|1 - \Pi\| \, \|x - \xi_h\|\end{aligned}$$

for any ξ_h in X_h. This proves (2.3). ∎

3. Application to Laplace, Lamé, and Stokes equations.

3.1. Laplace equation. The DPG method for the Laplace equation was first analyzed in [17] and the effect of inexact test spaces in a priori error analysis was clarified later in [22].

Given a Lipschitz polyhedron Ω in $\mathbb{R}^N$ and some $f \in L^2(\Omega)$, let u solve the Poisson problem

$$-\Delta u = f \text{ in } \Omega \quad \text{and} \quad u = 0 \text{ on } \partial\Omega. \tag{3.1}$$

Given a shape-regular triangulation Ω_h of Ω into simplices, set

$$\begin{aligned}
X &:= L^2(\Omega;\mathbb{R}^N) \times L^2(\Omega) \times H_0^{1/2}(\partial\Omega_h) \times H^{-1/2}(\partial\Omega_h),\\
Y &:= H(\operatorname{div},\Omega_h) \times H^1(\Omega_h), \text{ where}\\
H(\operatorname{div},\Omega_h) &:= \{\tau \in L^2(\Omega;\mathbb{R}^N) : \ \forall K \in \Omega_h,\ \tau|_K \in H(\operatorname{div},K)\},\\
H^1(\Omega_h) &:= \{v \in L^2(\Omega) : \ \forall K \in \Omega_h,\ v|_K \in H^1(K)\},\\
H_0^{1/2}(\partial\Omega_h) &:= \Big\{\eta \in \prod_{K\in\Omega_h} H^{1/2}(\partial K) : \exists\, w \in H_0^1(\Omega) \text{ such that}\\
&\qquad\qquad \forall K \in \Omega_h,\ \eta|_{\partial K} = w|_{\partial K}\Big\},\\
H^{-1/2}(\partial\Omega_h) &:= \Big\{\eta \in \prod_{K\in\Omega_h} H^{-1/2}(\partial K) : \exists\, q \in H(\operatorname{div},\Omega) \text{ such that}\\
&\qquad\qquad \forall K \in \Omega_h,\ \eta|_{\partial K} = q\cdot n|_{\partial K}\Big\}.
\end{aligned}$$

Throughout this paper, the L^2 scalar product $(\cdot,\cdot)_\Omega$ denotes the integral over Ω of the appropriate product of its arguments (which can be scalar, vector, or matrix-valued functions). In contrast, $(\cdot,\cdot)_{\Omega_h}$ abbreviates the sum of all scalar products $(\cdot,\cdot)_K$ over all $K \in \Omega_h$. The latter notation serves to clarify that any differential operators acting on its arguments are applied piecewise, element by element. The norms of $(v, q, \hat{u}, \hat{\sigma}_n)$

in $H^1(\Omega_h) \times H(\mathrm{div}, \Omega_h) \times H_0^{1/2}(\partial\Omega_h) \times H^{-1/2}(\partial\Omega_h)$ read

$$\begin{aligned}
\|v\|^2_{H^1(\Omega_h)} &= (v, v)_{\Omega_h} + (\mathrm{grad}\, v, \mathrm{grad}\, v)_{\Omega_h},\\
\|q\|^2_{H(\mathrm{div},\Omega_h)} &= (q, q)_{\Omega_h} + (\mathrm{div}\, q, \mathrm{div}\, q)_{\Omega_h},\\
\|\hat{u}\|_{H_0^{1/2}(\partial\Omega_h)} &= \inf_{w\in E_{\mathrm{grad}}(\hat{u})} \|w\|_{H^1(\Omega)},\\
\|\hat{\sigma}_n\|_{H^{-1/2}(\partial\Omega_h)} &= \inf_{q\in E_{\mathrm{div}}(\hat{\sigma}_n)} \|q\|_{H(\mathrm{div},\Omega)},\\
\|(\sigma, u, \hat{u}, \hat{\sigma}_n)\|^2_X &= \|\sigma\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)} + \|\hat{u}\|^2_{H^{1/2}(\partial\Omega_h)} + \|\hat{\sigma}_n\|^2_{H^{-1/2}(\partial\Omega_h)},\\
\|(\tau, v)\|^2_Y &= \|\tau\|^2_{H(\mathrm{div},\Omega_h)} + \|v\|^2_{H^1(\Omega_h)},
\end{aligned}$$

where $E_{\mathrm{grad}}(\hat{u}) = \{w \in H_0^1(\Omega) : \forall K \in \Omega_h,\ \hat{u}|_{\partial K} = w|_{\partial K}\}$ and $E_{\mathrm{div}}(\hat{\sigma}_n) = \{q \in H(\mathrm{div}, \Omega) : \forall K \in \Omega_h,\ \hat{\sigma}_n|_{\partial K} = q \cdot n|_{\partial K}\}$.

Reformulating (3.1) into the first order system $\sigma = -\mathrm{grad}\, u$, $\mathrm{div}\, \sigma = f$, an ultraweak formulation can be derived, which reads as follows: Given some f in $L^2(\Omega)$, find $x \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n) \in X$ satisfying

$$b(x, y) = F(y) \quad \text{for all } y \equiv (\tau, v) \in Y$$

with the forms $b(\cdot, \cdot)$ and $F(\cdot)$ defined by

$$\begin{aligned}
b(x, y) &:= (\sigma, \tau)_\Omega - (u, \mathrm{div}\, \tau)_{\Omega_h} + \langle \hat{u}, \tau \cdot n\rangle_{\partial\Omega_h} - (\sigma, \mathrm{grad}\, v)_{\Omega_h} + \langle \hat{\sigma}_n, v\rangle_{\partial\Omega_h},\\
F(y) &:= (f, v)_\Omega.
\end{aligned}$$

Here and throughout this paper, $\langle\cdot, \cdot\rangle_{\partial\Omega_h}$ denotes the sum of duality pairings which extend the appropriate L^2 inner product of scalar- or vector-valued functions on ∂K, over all $K \in \Omega_h$. Further details such as the formal equivalence of the ultraweak formulation with the strong and weak formulation of the Poisson model problem at hand, and a gentle derivation, can be found in [17].

Let $p \geq 0$ denote an integer, $P_p(K)$ denote the space of (algebraic) polynomials of degree at most p, with the convention that $P_{-1}(K)$ is trivial, $P_p(K; \mathbb{R}^N)$ denote the space of vector-valued functions with components in $P_p(K)$, and $R_p(K) := P_p(K, \mathbb{R}^N) + xP_p(K)$ denotes the standard Raviart–Thomas function space on a simplex K. Let $P_p(\Omega_h) = \{v : v|_K \in P_p(K), \forall K \in \Omega_h\}$. Let $\triangle_{N-1}(K)$ denote the set of all $(N-1)$-dimensional subsimplices of an N-simplex K. Define

$$\begin{aligned}
P_p(\partial K) &:= \{\mu \in L^2(\partial K) : \mu|_F \in P_p(F)\ \forall F \in \triangle_{N-1}(K)\},\\
\tilde{P}_p(\partial K) &:= P_p(\partial K) \cap C^0(\partial K).
\end{aligned}$$

Set the finite-dimensional trial space X_h and the test space Y_h by

$$\begin{aligned}
&X_h := \{(\sigma, u, \hat{u}, \hat{\sigma}_n) \in X : \forall K \in \Omega_h,\ \sigma|_K \in P_p(K; \mathbb{R}^N),\ u|_K \in P_p(K),\\
&\qquad\qquad \hat{u}|_{\partial K} \in \tilde{P}_{p+1}(\partial K),\ \hat{\sigma}_n|_{\partial K} \in P_p(\partial K)\}, \qquad (3.2a)\\
&Y_h := \{(\tau, v) \in Y : \forall K \in \Omega_h,\ \tau|_K \in R_{p+1}(K),\ v|_K \in P_{p+N}(K)\}. \qquad (3.2b)
\end{aligned}$$

This example meets all the assumptions of the abstract a posteriori analysis of the previous section, as seen next. Let f_K denote the integral mean of f over a simplex K of diameter h_K.

THEOREM 3.1. *There exist mesh-size independent positive constants C_1, C_2, and C_3 such that the exact solution x of the ultraweak formulation for the Laplace equation* *and* any *discrete function* $\tilde{x}_h \in X_h$ *with its error estimator* $\tilde{\eta} := \|F - B\tilde{x}_h\|_{Y_h^*}$ *satisfy*

$$C_1\|x - \tilde{x}_h\|_X^2 - C_2 \operatorname{osc}(F)^2 \le \tilde{\eta}^2 \le C_3\|x - \tilde{x}_h\|_X^2.$$

Furthermore, if $p \ge 1$*, there is a mesh-independent constant* $C_4 > 0$ *such that*

$$\operatorname{osc}(F)^2 \le C_4 \sum_{K\in\Omega_h} \|h_K(f - f_K)\|_{L^2(K)}^2. \tag{3.3}$$

The remaining parts of this section provide some details on (H1)–(H3) so that Theorem 3.1 is identified as a particular case of Theorem 2.1. Let $0 < C < \infty$ denote a generic constant, whose values at different occurrences may differ, but will remain independent of h_K for all $K \in \Omega_h$. The value of C may depend on the shape (angles) of the simplices and on p.

LEMMA 3.2 (see [22]). *For any* $p \ge 0$*, there exists a* C *and a bounded linear operator* $\Pi_{p+N}^{\mathrm{grad}} : H^1(K) \to P_{p+N}(K)$ *such that for all* $v \in H^1(K)$*,*

$$(\Pi_{p+N}^{\mathrm{grad}} v - v, q_{p-1})_K = 0, \qquad \textit{for all } q_{p-1} \in P_{p-1}(K), \tag{3.4a}$$

$$\langle \Pi_{p+N}^{\mathrm{grad}} v - v, \mu_p\rangle_{\partial K} = 0, \qquad \textit{for all } \mu_p \in P_p(\partial K), \tag{3.4b}$$

$$\|\Pi_{p+N}^{\mathrm{grad}} v\|_{H^1(K)} \le C\|v\|_{H^1(K)}, \tag{3.4c}$$

$$\|v - \Pi_{p+N}^{\mathrm{grad}} v\|_{L^2(K)} \le C h_K |v|_{H^1(K)}. \tag{3.4d}$$

Proof. Properties (3.4a)–(3.4c) can be found in [22, Lemma 3.2]. Although (3.4d) is not explicitly stated in [22], it follows from the arguments therein: The design of $\Pi_{p+N}^{\mathrm{grad}}$ in [22] shows $v - \Pi_{p+N}^{\mathrm{grad}} v = \Pi_0(v - v_K)$ for the integral mean v_K of v over K and some bounded linear operator $\Pi_0 : H^1(K) \to L^2(K)$. Then (3.4d) follows from a Poincaré inequality. ∎

LEMMA 3.3 (see [22, Lemma 3.3]). *For any* $p \ge 0$*, there exists a* C *and a bounded linear operator* $\Pi_{p+2}^{\mathrm{div}} : H(\mathrm{div}, K) \mapsto P_{p+2}(K;\mathbb{R}^N)$ *such that for all* $\tau \in H(\mathrm{div}, K)$*,*

$$(\Pi_{p+2}^{\mathrm{div}}\tau - \tau, q_p)_K = 0, \qquad \textit{for all } q_p \in P_p(K;\mathbb{R}^N), \tag{3.5a}$$

$$\langle(\Pi_{p+2}^{\mathrm{div}}\tau - \tau)\cdot n, \mu_{p+1}\rangle_{\partial K} = 0 \qquad \textit{for all } \mu_{p+1} \in \tilde{P}_{p+1}(\partial K), \tag{3.5b}$$

$$\|\Pi_{p+2}^{\mathrm{div}}\tau\|_{H(\mathrm{div},K)} \le C\|\tau\|_{H(\mathrm{div},K)}. \tag{3.5c}$$

Proof of Theorem 3.1. The boundedness of the bilinear form b in (H1) is shown in [17, section 4.4]. Recall from [17, Theorem 4.2 and Lemma 4.1] that

$$\{x \in X :\ b(x, y) = 0\ \forall y \in Y\} = \{0\}$$

and that

$$0 < \beta' := \inf_{y\in Y\setminus\{0\}} \sup_{x\in X\setminus\{0\}} \frac{b(x, y)}{\|x\|_X\,\|y\|_Y}$$

with some mesh-size independent constant β'. By duality of linear operators, this implies (H2) with $\beta = \beta'$.

From Lemmas 3.2 and 3.3, it is easy to see that the tensor product Π,

$$\Pi(\tau, v) := (\Pi_{p+2}^{\mathrm{div}}\tau, \Pi_{p+N}^{\mathrm{grad}} v) \quad \text{for all } (\tau, v) \in Y,$$

verifies (H3). Hence the reliability and efficiency in Theorem 3.1 follow from Theorem 2.1.

Moreover, the data approximation term $\mathrm{osc}(F)$ from (1.12) can be simplified: For any $y \equiv (\tau, v) \in Y$ and any $f_{p-1} \in P_{p-1}(\Omega_h)$,

$$\begin{aligned}(F \circ (1 - \Pi))(y) = F(y - \Pi y) &= (f, v - \Pi^{\mathrm{grad}}_{p+N} v)_{\Omega_h} \\ &= (f - f_{p-1}, v - \Pi^{\mathrm{grad}}_{p+N} v)_{\Omega_h}.\end{aligned}$$

Thus (3.4d) implies (3.3). $\square$

Similar arguments also prove the reliability and efficiency of the estimator for a different DPG method for the Laplacian in [18], an example omitted for brevity.

3.2. Linearized elasticity. The DPG method of [7] for a linear elastic body Ω in $N = 2$ or 3 space dimensions concerns the boundary value problem which, given some volume force $f : \Omega \to \mathbb{R}^N$, seeks the displacement $u : \Omega \to \mathbb{R}^N$ and stress $\sigma : \Omega \to \mathbb{S}$ satisfying

$$\begin{aligned} A\sigma &= \varepsilon(u) && \text{on } \Omega, \\ \operatorname{div} \sigma &= f && \text{on } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{3.6}$$

Here and throughout this paper, $\mathbb{S} = \mathrm{sym}(\mathbb{M})$ and $\mathbb{K} = \mathrm{skw}(\mathbb{M})$, where $\mathrm{sym}(M) = (M + M^T)/2$ and $\mathrm{skw}(M) = (M - M^T)/2$ for any matrix $M \in \mathbb{M} := \mathbb{R}^{N\times N}$. The inner product between matrices σ and τ is the Frobenius inner product $\sigma : \tau = \mathrm{tr}(\sigma' \tau)$ with the trace $\mathrm{tr}\,\tau := \tau_{11} + \cdots + \tau_{NN} = \tau : I_{N\times N}$. The divergence on matrix fields is taken rowwise and $\varepsilon(u) = \mathrm{sym}(\mathrm{grad}\, u)$ is the linear Green strain. Let the trace-free deviatoric part of any matrix $\tau \in \mathbb{M}$ be denoted by $\mathrm{dev}\,\tau = \tau - \frac{1}{N}\,\mathrm{tr}(\tau) I_{N\times N}$. The compliance tensor $A \equiv A(x)$ is a self-adjoint positive definite fourth-order tensor, which (for simplicity) is isotropic in that the positive Lamé parameters $P, Q \in P_0(\Omega_h)$ satisfy

$$A\tau = P\ \mathrm{dev}\,\tau + \frac{Q}{N}\ \mathrm{tr}(\tau) I_{N\times N} \qquad \text{for all } \tau \in \mathbb{M}.$$

The weak formulation of linear elasticity is included in the abstract setting for

$$X = L^2(\Omega; \mathbb{M}) \times L^2(\Omega; \mathbb{R}^N) \times H_0^{1/2}(\partial\Omega_h; \mathbb{R}^N) \times H^{-1/2}(\partial\Omega_h; \mathbb{R}^N) \times \mathbb{R}, \tag{3.7a}$$

$$Y = H(\mathrm{div}, \Omega_h; \mathbb{S}) \times H^1(\Omega_h; \mathbb{R}^N) \times L^2(\Omega; \mathbb{K}) \times \mathbb{R}. \tag{3.7b}$$

Here and throughout this paper, notation for spaces of scalar-valued functions is extended to analogous spaces of vector-valued and matrix-valued functions without much ado. For instance, $L^2(\Omega; \mathbb{M})$ denotes the set of matrix-valued functions whose entries belong to $L^2(\Omega)$; $H_0^{1/2}(\partial\Omega_h; \mathbb{R}^N)$ consists of vector functions with components in $H_0^{1/2}(\partial\Omega_h)$; $H(\mathrm{div}, \Omega_h; \mathbb{S})$ denotes symmetric matrix-valued functions with rows in $H(\mathrm{div}, \Omega_h)$, etc. The notation for polynomial spaces is extended similarly, e.g., symmetric, skew-symmetric, and general matrix-valued functions with entries in $P_p(K)$ read $P_p(K; \mathbb{S})$, $P_p(K; \mathbb{K})$, and $P_p(K; \mathbb{M})$.

The weak formulation of (3.6) of [7] seeks $x \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in X$ with

$$b(x, y) = F(y) \qquad \text{for all } y \equiv (\tau, v, q, \omega) \in Y,$$

where Q_0 is the essential infimum of $Q(x)$ over $x \in \Omega$ and

$$\begin{aligned} b(x, y) &:= (A\sigma, \tau)_\Omega + (u, \operatorname{div} \tau)_{\Omega_h} - \langle \hat{u}, \tau\, n\rangle_{\partial\Omega_h} + Q_0^{-1}(\alpha I_{N\times N}, A\tau)_\Omega \\ &\quad + (\sigma, \nabla v)_{\Omega_h} + (\sigma, q)_\Omega - \langle \hat{\sigma}_n, v\rangle_{\partial\Omega_h} + Q_0^{-1}(A\sigma, \omega I_{N\times N})_\Omega, \\ F(y) &:= -(f, v)_\Omega. \end{aligned} \tag{3.7c}$$

The DPG solution x_h for linear elasticity is defined by (1.3) for $p \in \mathbb{N}_0$ and

$$X_h := \{(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in X : \forall K \in \Omega_h,\ \sigma|_K \in P_p(K;\mathbb{M}), u|_K \in P_p(K;\mathbb{R}^N), \hat{u}|_{\partial K} \in \tilde{P}_{p+1}(\partial K;\mathbb{R}^N),\ \hat{\sigma}_n|_{\partial K} \in P_p(\partial K;\mathbb{R}^N), \alpha \in \mathbb{R}\}, \tag{3.7d}$$

$$Y_h := \{(\tau, v, q, \omega) \in Y : \forall K \in \Omega_h,\ \tau|_K \in P_{p+2}(K;\mathbb{S}), v|_K \in P_{p+N}(K;\mathbb{R}^N),\ q|_K \in P_p(K;\mathbb{K}),\ \omega \in \mathbb{R}\}. \tag{3.7e}$$

The second part of this section provides some details on (H1)–(H3) such that Theorem 2.1 guarantees reliable and efficient error control. The verification of (H3) utilizes the following operator constructed in [22, Lemma 4.1] using the degrees of freedom for symmetric matrices with polynomial entries given in [21]. An operator different from the tensor product (matrix version) of the operator Π^{div}_{p+2} given in Lemma 3.3 is required due to the symmetry of τ.

LEMMA 3.4 (see [22, Lemma 4.1]). *For any $p \geq 0$, there exists a $C > 0$ and a operator $\Pi^{\mathrm{div},\mathbb{S}}_{p+2} : H(\mathrm{div}, K;\mathbb{S}) \to P_{p+2}(K;\mathbb{S})$ such that every $\tau \in H(\mathrm{div}, K;\mathbb{S})$ satisfies*

$$(\Pi^{\mathrm{div},\mathbb{S}}_{p+2}\tau, q_p)_K = (\tau, q_p)_K \qquad \textit{for all } q_p \in P_p(K;\mathbb{S}), \tag{3.8a}$$

$$\langle \Pi^{\mathrm{div},\mathbb{S}}_{p+2}\tau \cdot n, \mu_{p+1}\rangle_{\partial K} = \langle \mu_{p+1}, \tau \cdot n\rangle_{1/2,\partial K} \qquad \textit{for all } \mu_{p+1} \in \tilde{P}_{p+1}(\partial K;\mathbb{R}^N), \tag{3.8b}$$

$$\|\Pi^{\mathrm{div},\mathbb{S}}_{p+2}\tau\|_{H(\mathrm{div},K)} \leq C\|\tau\|_{H(\mathrm{div},K)}. \tag{3.8c}$$

To verify (H3), apply the operator of Lemma (3.4) to τ, the operator of Lemma 3.2 to each component of v, and the L^2-orthogonal projection $P^{\mathbb{K}}_p q$ of q into $P_p(K,\mathbb{K})$, i.e., set

$$\Pi y = (\Pi^{\mathrm{div},\mathbb{S}}_{p+2}\tau, \Pi^{\mathrm{grad}}_{p+N}v, P^{\mathbb{K}}_p q, \omega)$$

for all $y \equiv (\tau, v, q, \omega)$ in Y.

LEMMA 3.5 (see [7, Lemma 5]). *There exists a constant $C_0 > 0$ independent of Ω_h and Q such that* (H2) *holds for the DPG method* (3.7) *with*

$$\beta = C_0/\zeta^4 \quad \textit{and} \quad \zeta = \|Q\|_{L^\infty(\Omega)}/Q_0. \quad \square$$

This is a *locking-free* inf-sup condition because β does not degenerate when a homogeneous material approaches the incompressibility limit (i.e., when Q is constant on Ω and $Q \to 0$). The aforementioned results allow immediate verification of assumptions (H1)–(H3) so that Theorem 2.1 leads to the reliability and efficiency of the error estimator for the linear elasticity formulation (3.7). The constants in these estimates are independent of Q when Q is constant on Ω.

The remaining parts of this section concern the analysis of a symmetric variant of the above DPG method, which strongly imposes symmetry of the stress tensor and seeks σ in $L^2(\Omega;\mathbb{S})$. This variant, although not presented in [7], can be easily analyzed using the above results. Given X, Y, X_h, Y_h, b as above in (3.7), ignore the variable q from the previous analysis and set

$$X_2 := \{(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in X : \sigma^T = \sigma\} \quad \text{and} \quad X_{2,h} := X_h \cap X_2, \tag{3.9a}$$

$$Y_2 := H(\mathrm{div}, \Omega_h;\mathbb{S}) \times H^1(\Omega_h;\mathbb{R}^N) \times \mathbb{R} \tag{3.9b}$$

$$Y_{2,h} := \{(\tau, v, \omega) \in Y_2 :\ (\tau, v, 0, \omega) \in Y_h\}, \tag{3.9c}$$

$$b_2((\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, \omega)) = b((\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, 0, \omega)). \tag{3.9d}$$

The new symmetric variant of the DPG method for linear elasticity seeks $x_2 \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha)$ in X_2 with

$$b_2(x_2, y_2) = F(y_2) \qquad \text{for all } y_2 \equiv (\tau, v, \omega) \in Y_2.$$

THEOREM 3.6. *The statements of Theorem* 3.1 *hold for the DPG formulations* (3.7) *and* (3.9) *in linear elasticity.*

Proof. The previous analysis of this section establishes (H1)–(H3) for the formulation (3.7). Assumptions (H1) and (H3) for the formulation (3.9) follow just as for (3.7). Assumption (H2) is verified for the symmetric formulation (3.9) in the remaining part of this proof with $\beta = C_0/\zeta^4$ from Lemma 3.5.

Since $(\sigma, q)_{\Omega_h} = 0$ for all $\sigma \in X_2$ and all $q \in L^2(\Omega; \mathbb{K})$, Lemma 3.5 implies, for all $x_2 \in X_2$, that

$$\beta \|x_2\|_X \le \sup_{(\tau,v,q,\omega)\in Y\setminus\{0\}} \frac{b(x_2, (\tau, v, q, \omega))}{\|(\tau, v, q, \omega)\|_Y} \le \sup_{(\tau,v,\omega)\in Y_2\setminus\{0\}} \frac{b_2(x_2, (\tau, v, \omega))}{\|(\tau, v, \omega)\|_{Y_2}}.$$

To complete the verification of (H2), it remains to prove that any $y_2 \equiv (\tau, v, \omega) \in Y_2$ satisfying $b_2(x_2, y_2) = 0$ for all $x_2 \in X_2$ must vanish. To prove this, observe

$$b((\sigma + \eta, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, q, \omega)) = b_2(x_2, y_2) - (r, \operatorname{skw}(\operatorname{grad} v) + q)_\Omega$$

for all $x_2 \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha)$ in X_2 and all $r, q \in L^2(\Omega; \mathbb{K})$. Given $y_2 \equiv (\tau, v, \omega) \in Y_2$ with $b_2(x_2, y_2) = 0$ for all $x_2 \in X_2$ define $q := -\operatorname{skw}(\operatorname{grad} v)$. The right-hand side in the aforementioned identity vanishes. Consequently, $y := (\tau, v, q, \omega)$ satisfies $b(x, (\tau, v, q, \omega)) = 0$ for all $x \in X$. Since (H2) holds for (3.7), $(\tau, v, q, \omega) = 0$. ∎

3.3. A new DPG formulation for Stokes flow. The equations of Stokes describe the steady state velocity $u : \Omega \to \mathbb{R}^N$ and pressure $p : \Omega \to \mathbb{R}$ of a homogeneous fluid of constant viscosity $\nu = 1/2$ via

$$-\nu \Delta u + \operatorname{grad} p = -f \qquad \text{in } \Omega, \tag{3.10a}$$

$$\operatorname{div} u = 0 \qquad \text{in } \Omega. \tag{3.10b}$$

This is complemented by the simple homogeneous boundary condition $u = 0$ on $\partial\Omega$ and the side restriction $(p, 1)_\Omega = 0$. The pseudostress $\tilde{\sigma} := 1/2 \operatorname{grad} u - p\delta$ is often used to reformulate (3.10) into a first order system [8, 12, 30], but $\tilde{\sigma}$ is not symmetric in general and is not the true fluid stress.

Within the DPG formalism however, it is easy to maintain symmetry. Hence a first order formulation with the symmetric physical stress $\sigma = \varepsilon(u) - pI_{N\times N}$ serves as the starting point:

$$\begin{aligned} \sigma + p\, I_{N\times N} - \varepsilon(u) &= 0 \quad \text{in } \Omega, \\ \operatorname{div} \sigma &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= 0 \quad \text{in } \Omega. \end{aligned}$$

Using the notation of the previous subsection for the trace-free deviatoric part, these equations give an equivalent *stress-velocity* reformulation:

$$\operatorname{dev} \sigma - \varepsilon(u) = 0 \text{ in } \Omega, \tag{3.11a}$$

$$\operatorname{div} \sigma = f \text{ in } \Omega, \tag{3.11b}$$

$$u = 0 \text{ on } \partial\Omega \quad \text{and} \quad \int_\Omega \operatorname{tr} \sigma \, dx = 0. \tag{3.11c}$$

Given any element domain $K \in \Omega_h$, test the first two equations of (3.11) with smooth

$\tau : K \to \mathbb{S}$ and $v : K \to \mathbb{R}^N$ and integrate by parts. The sum over all $K \in \Omega_h$ reads

$$\begin{aligned}(\operatorname{dev}\sigma, \tau)_\Omega + (u, \operatorname{div}\tau)_{\Omega_h} - \langle \hat{u}, \tau n\rangle_{\partial\Omega_h} &= 0,\\ -(\sigma, \varepsilon(v))_{\Omega_h} + \langle \hat{\sigma}_n, v\rangle_{\partial\Omega_h} &= (f, v)_\Omega,\end{aligned}$$

where the traces of σn and u on element interfaces are set as independent unknowns $\hat{\sigma}_n$ and $\hat{u}$. Using a Lagrange multiplier α to model the constraint (3.11c), the variational problem is to find $x \equiv (\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha) \in X_2$ such that $b(x, y) = F(y)$ for all $y \equiv (\tau, v, \omega) \in Y_2$ with $X_2, X_{2,h}, Y_2, Y_{2,h}$ from (3.9a)–(3.9c) and

$$\begin{aligned}b(x,y) &:= (\operatorname{dev}\sigma, \tau)_\Omega + (u, \operatorname{div}\tau)_{\Omega_h} - \langle \hat{u}, \tau n\rangle_{\partial\Omega_h} + (\alpha, \operatorname{tr}\tau)_\Omega\\ &\quad + (\sigma, \varepsilon(v))_{\Omega_h} - \langle \hat{\sigma}_n, v\rangle_{\partial\Omega_h} + (\operatorname{tr}\sigma, \omega)_\Omega,\\ F(y) &:= -(f, v)_\Omega.\end{aligned} \tag{3.12}$$

This novel DPG method admits the following a priori and a posteriori error estimates for the x_h from (1.3) and the η from (1.8). (Of course, unlike the a priori estimate, the a posteriori estimate holds even if x_h is replaced by any $\tilde{x}_h \in X_{2,h}$ for the reasons amply clarified previously.)

THEOREM 3.7. *For any $p \geq 0$, there exist mesh-independent constants $C_1, \ldots, C_4 > 0$ such that*

$$\|x - x_h\|_X \leq C_1 \min_{\xi_h \in X_h} \|x - \xi_h\|_X, \tag{3.13}$$

$$C_4\|x - x_h\|_X^2 - C_2 \operatorname{osc}(F)^2 \leq \eta^2 \leq C_3\|x - x_h\|_X^2. \tag{3.14}$$

If $p \geq 1$ then $\operatorname{osc}(F)$ satisfies (3.3).

Proof. Since the verification of (H1) is immediate, this proof focuses on the proof of (H2) and (H3). For all $y \equiv (\tau, v, \omega)$ in Y set

$$\Pi y = (\Pi_{p+2}^{\operatorname{div},\mathbb{S}}\tau, \Pi_{p+N}^{\operatorname{grad}} v, \omega) \in Y_h,$$

where the operator $\Pi_{p+N}^{\operatorname{grad}}$ of Lemma 3.2 applies componentwise to the vector valued v and the operator $\Pi_{p+2}^{\operatorname{div},\mathbb{S}}$ from Lemma 3.4 applies to the symmetric matrix function τ. The properties of Π verify assumption (H3).

The proof of (H2) utilizes a perturbation argument that links the Stokes formulation to the elasticity formulation (3.9). Let the bilinear form $b(\cdot,\cdot)$ in (3.12) be denoted by $b_0(\cdot,\cdot)$ for the remaining parts of this proof. Given $0 < \kappa < 1$, define the compliance tensor A as in linear elasticity with constant $P \equiv 1$ and $Q \equiv \kappa$ and denote the resulting elasticity form $b(\cdot,\cdot)$ of (3.9d) by $b_\kappa(\cdot,\cdot)$. From the definitions,

$$b_\kappa((\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, \omega)) = b_0(\sigma, u, \hat{u}, \hat{\sigma}_n, \alpha), (\tau, v, \omega)) + \frac{\kappa}{N}(\operatorname{tr}\sigma, \operatorname{tr}\tau)_\Omega.$$

The linear operators $B_\kappa : X \to Y^*$ associated with $b_\kappa(\cdot,\cdot)$ satisfy

$$\|B_0 - B_\kappa\| \leq c\kappa \tag{3.15}$$

with some κ-independent constant $c > 0$.

By the previous choice of A, Lemma 3.5 implies that $\beta = C_0$, a constant independent of κ. Hence B_κ is invertible and $\|B_\kappa^{-1}\| \leq C_0^{-1}$. This and the formula

$$B_{\kappa_1}^{-1} - B_{\kappa_2}^{-1} = B_{\kappa_1}^{-1}(B_{\kappa_2} - B_{\kappa_1})B_{\kappa_2}^{-1}$$

imply that the operator $B \equiv B_0$ is a bijection and $\|B^{-1}\| \leq C_0^{-1}$. This completes the verification of (H2). Hence (3.14) follows from Theorem 2.1 and (3.13) follows from Remark 2.5. $\square$

4. Numerical illustrations. Results using adaptive and uniform h-refinements of triangular meshes (on two-dimensional domains) for the previously considered boundary value problems are reported in this section. (No p-refinements are considered.) The Y-norm of the restriction of ε_h on any mesh element $K \in \Omega_h$ serves as a refinement indicator that marks elements for refinement in an adaptive loop. In each adaptive iteration, h-refinement is performed via a 50% maximum criterion, the discrete solution x_h is computed on the current mesh, the corresponding error representation ε_h is locally computed, and further elements are marked for refinement as necessary based on the elementwise norms of ε_h.

All adaptive refinements in the reports below use 1-irregular triangular meshes (with at most one hanging node per edge). These meshes pose no difficulty in the presented theory if one uses the so-called minimum rule: When a parent mesh edge is split into two child edges by a hanging node, the minimum rule sets interface variables to polynomials (not piecewise polynomials) on the parent edge, so definitions for trial and test spaces like (3.2a)–(3.2b) and (3.7d)–(3.7e) remain unchanged. Since Π was constructed element by element in all the previous examples, verification of (H3) on 1-irregular meshes is immediate for all those examples. Note that verification of (H1) and (H2) never used mesh conformity.

Denote the effectivity index for the discrete solution x_h and its perturbations $\tilde{x}_h$ by

$$\rho = \frac{\eta}{\|x - x_h\|_X}, \qquad \tilde{\rho} = \frac{\tilde{\eta}}{\|x - \tilde{x}_h\|_X},$$

where η and $\tilde{\eta}$ are defined previously and locally computable—see (1.6) and Theorem 2.1. Although the constants in the estimates of Theorem 2.1 have been shown to be mesh independent for the previously considered examples, their values influence the effectivity indices. In particular, the theory does not guarantee that the values of ρ and $\tilde{\rho}$ are less than one. However, the numerical reports below show that their values are fairly close to one in the examples considered.

4.1. The Laplace example. Consider the method of section 3.1 that yields an approximation to $x = (\sigma, u, \hat{u}, \hat{\sigma}_n)$, namely, $x_h = (\sigma_h, u_h, \hat{u}_h, \hat{\sigma}_{n,h})$ in the space X_h defined in (3.2a). Even in cases where the exact solution x is known, the computation of $\|x - x_h\|_X$ is nontrivial because the components $\|\hat{u} - \hat{u}_h\|_{H_0^{1/2}(\partial\Omega_h)}$ and $\|\hat{\sigma} - \hat{\sigma}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)}$ in the norm are defined using an infimum over an infinite-dimensional hyperspace. Here and in the remainder, such quotient norms are always approximated by an infimum over an appropriate finite-dimensional subset, e.g.,

$$\|\hat{u} - \hat{u}_h\|_{H_0^{1/2}(\partial\Omega_h)} = \inf_{w \in E_{\mathrm{grad}}(\hat{u} - \hat{u}_h)} \|w\|_{H^1(\Omega)}$$

is approximated by

$$\inf_{w_h \in E^h_{\mathrm{grad}}(\hat{u}_I - \hat{u}_h)} \|w_h\|_{H^1(\Omega)},$$

a computable infimum over the finite-dimensional hyperspace $E^h_{\mathrm{grad}}(\hat{u}_I - \hat{u}_h) = \{w_h \in H_0^1(\Omega) :\ \forall K \in \Omega_h,\ w|_{\partial K} = (u_I - \hat{u}_h)|_{\partial K} \text{ and } w_h|_K \in P_{p+1}(K)\}$ where u_I is an H^1-interpolant of u of degree $p+1$. Then the next higher order element bubble functions are added and the minimum is recomputed. The process is repeated until the result (to a few significant digits) of the minimization remains unchanged. Similarly,

$$\|\hat{\sigma}_n - \hat{\sigma}_{n,h}\|_{H^{-1/2}(\partial\Omega_h)} = \inf_{q \in E_{\mathrm{div}}(\hat{\sigma}_n - \hat{\sigma}_{n,h})} \|q\|_{H(\mathrm{div},\Omega)}$$

is approximated by the infimum of $\|q_h\|_{H(\mathrm{div},\Omega)}$ over all q_h in a Raviart–Thomas subspace of $H(\mathrm{div},\Omega)$ such that $q_h \cdot n|_{\partial K} = (\sigma_I \cdot n - \hat{\sigma}_{n,h})|_{\partial K}$ for all $K \in \Omega_h$, where σ_I is an interpolant of σ into the Raviart–Thomas space. These approximations, or its appropriate modifications for each boundary value problem, are used in each report below. At times the simplifications afforded by the problem are exploited, e.g., if $\hat{u}$ is the trace of a harmonic solution, the norm in the minimization is modified to an equivalent graph norm to avoid interpolation of exact solution.

The first experiment solves the DPG method for the Laplace equation on the unit square with load set so that the exact solution is $u = \sin(\pi x_1)\sin(\pi x_2)$. For a series of successive uniform h-refinements, the error in computed solution and the corresponding effectivity indices ρ are shown in Figures 1(a) and 1(c) for $p = 0, 1,$ and 2. Noting that the number of degrees of freedom N in the uniform refinement case is $O(h^{-2})$, the observed rates of convergence are as expected from the theory. The effectivity indices are fairly close to one, indicating that the error estimator and error are close.

Next, consider the Dirichlet problem on the L-shaped domain $\Omega = (-1,1) \times (-1,1) \backslash [0,1] \times [-1,0]$ with boundary conditions set to yield the singular exact solution $u(r,\theta) = r^{2/3} \sin\left(\frac{2}{3}(\theta + \frac{\pi}{2})\right)$. Results from an h-adaptive algorithm using the DPG method with $p = 2$ are reported in Figure 1(b). The error estimator is observed to closely follow the actual error. Furthermore, near the end of adaptive iteration, despite the singularity in the solution, the rate of convergence appears to stabilize to the same rate as in the uniformly refined smooth solution case of Figure 1(a) ($p = 2$).

The remaining plots in the panel simulate the effects of inexact solvers, or partially converged iterative solutions. This is done by randomly perturbing the values of the degrees of freedom of the computed solutions by 5% using a pseudorandom number generator. The effectivity indices $\tilde{\rho}$ obtained when perturbing the solutions in uniformly and adaptively refined cases are shown in Figures 1(e) and 1(f), respectively. The values remain close to one and show that the error estimator remains useful for inexact solutions. This completes the numerical illustration of the estimates of Theorem 3.1 for both the DPG solution x_h and an arbitrary $\tilde{x}_h$ in the Laplace example.

4.2. The elasticity example. Experiments similar to section 4.1 were also performed with the formulation (3.9), both on the unit square and the L-shaped domain. On the unit square, setting unit Lamé parameters, and the load so that both components of the exact displacement equal $\sin(\pi x_1)\sin(\pi x_2)$, plots of quantities analogous to those in Figure 1 are reported in Figure 2. On the L-shaped domain, the planar elastic properties are set to that of steel and an h-adaptive algorithm is run on the DPG method with $p = 2$. More details on the material properties of steel and the corresponding singular solution with a stress singularity of strength $r^{-0.3962\ldots}$ can be found in [7, section 7.2.2]. Effectivity is reported both for the computed solution and for the 5% random perturbations, as before. In all cases, the values of indices ρ and $\tilde{\rho}$ are comparable to the corresponding quantities in Figure 1. Further numerical results for the other DPG formulation for linear elasticity (3.7), including results from hp-adaptivity, have already been presented in [7].

4.3. The Stokes example. The Stokes problem is solved using the new DPG method of section 3.3 on the unit square with data set so that the exact solution is $u = \mathrm{curl}(\sin(\pi x_1)\sin(\pi x_2))$. The h-convergence rates for this new method have not been previously reported. The first plot in Figure 3 shows these rates for $p = 0, 1,$ and 2, as the unit square is successively refined in a uniform fashion. Clearly, the number of degrees of freedom N is $O(h^{-2})$ in these uniform refinements, so these observed

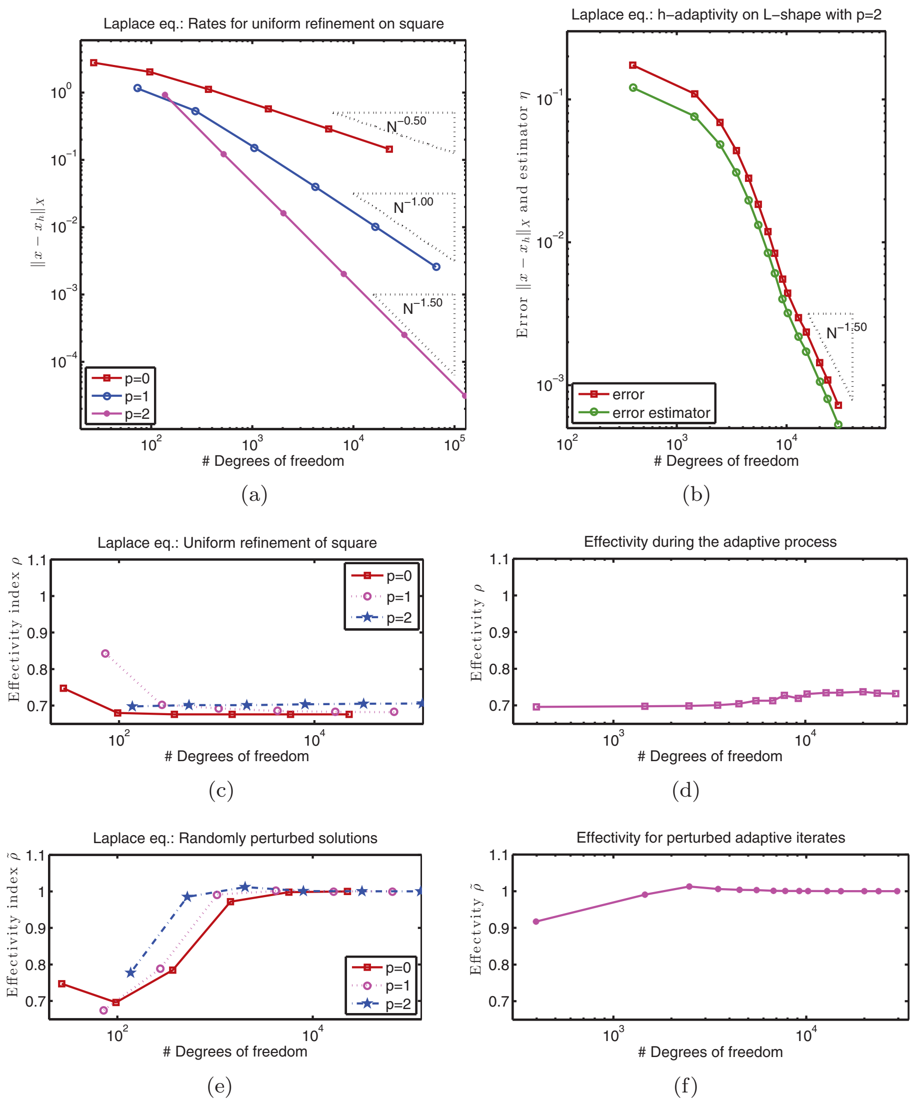

FIG. 1. *Results from the Laplace example for a smooth solution (left) and a singular solution (right).* (a) *Rates of error reduction for a smooth solution under uniform refinements of the unit square.* (b) *Error and estimator in an h-adaptive algorithm for a singular solution with $p = 2$ on an L-shaped domain.* (c) *Effectivity indices for the solutions of* (a). (d) *Effectivity indices for the solutions of* (b). (e) *Changes in effectivity when solutions of* (a) *are perturbed at each uniform refinement.* (f) *Changes in effectivity when solutions of* (b) *are perturbed at each adaptive refinement.*

rates are as predicted by the a priori estimate (3.13). The a posteriori estimate (3.14) is illustrated by the plot of ρ. The remaining plot of $\tilde{\rho}$ shows that the estimator remains effective even when the solution computed at each refinement is randomly perturbed by 5% as described for the previous examples.

Results from an h-adaptive algorithm applied to the DPG method for Stokes flow on an L-shaped domain are also portrayed in Figure 3. To describe this well-known [14, 23, 29] singular Stokes solution of Osborn, let $z = 0.544484\ldots$ be the

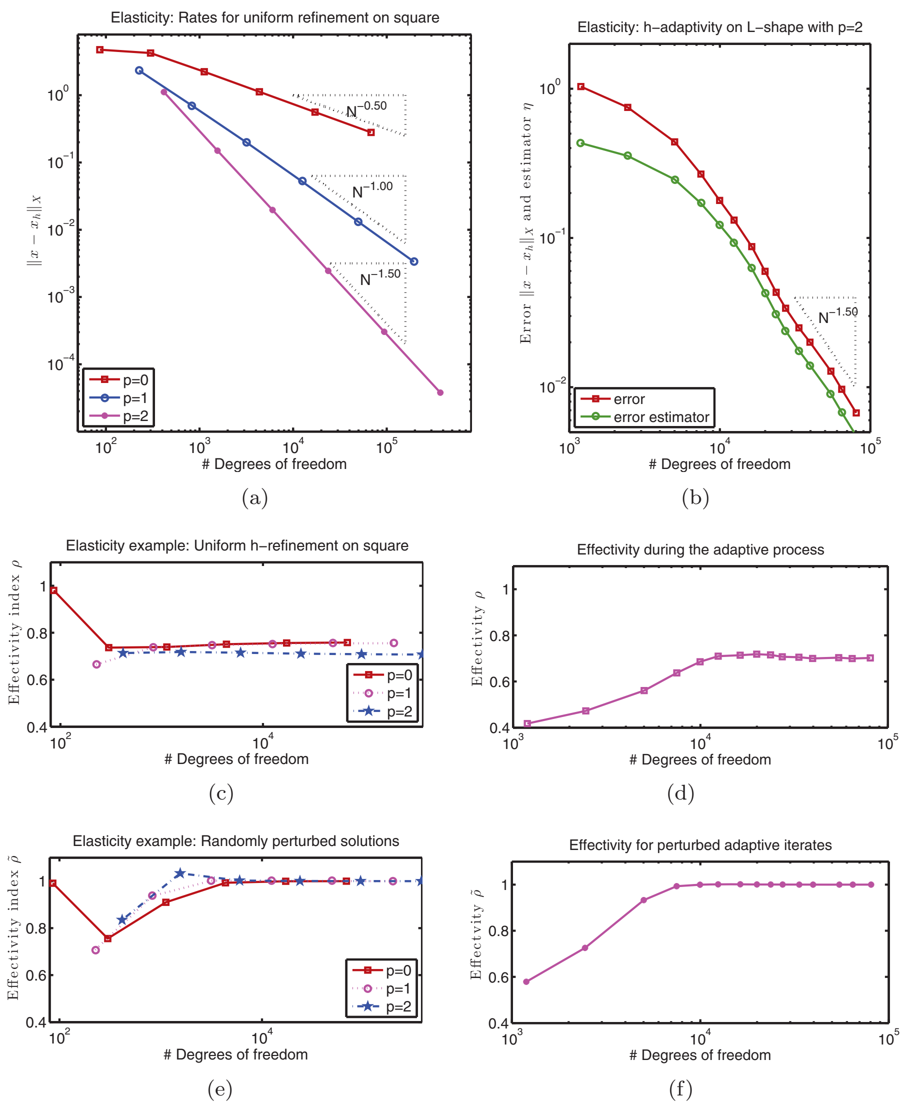

FIG. 2. *Results from the elasticity example for a smooth solution (left) and a singular solution (right).* (a) *Convergence rates for a smooth solution under uniform refinements of the unit square.* (b) *Error and estimator in an h-adaptive algorithm with $p = 2$ on L-shaped steel.* (c) *Effectivity indices for the solutions of* (a). (d) *Effectivity indices for the solutions of* (b). (e) *Changes in effectivity when solutions of* (a) *are perturbed at each uniform refinement.* (f) *Changes in effectivity when solutions of* (b) *are perturbed at each adaptive refinement.*

nontrivial solution of $z^2 = \sin^2(3z\pi/2)$ with the smallest positive real part, and let $s_\pm = r^{1+z}\sin((z \pm 1)\theta)$, $c_\pm = r^{1+z}\cos((z \pm 1)\theta)$, and $a_\pm = -z\cot(3z\pi/2)/(z \pm 1)$. The singular Stokes solution with $u = 0$ on the two edges meeting the origin is $u = \operatorname{curl}(a_+ s_+ + a_- s_- + c_+ - c_-)$. On all except those two edges, the no-slip boundary condition of (3.11) was modified to traction boundary conditions and the correspondingly revised DPG method was implemented. The results in Figure 3 show the utility of the estimator in capturing the singularity correctly.

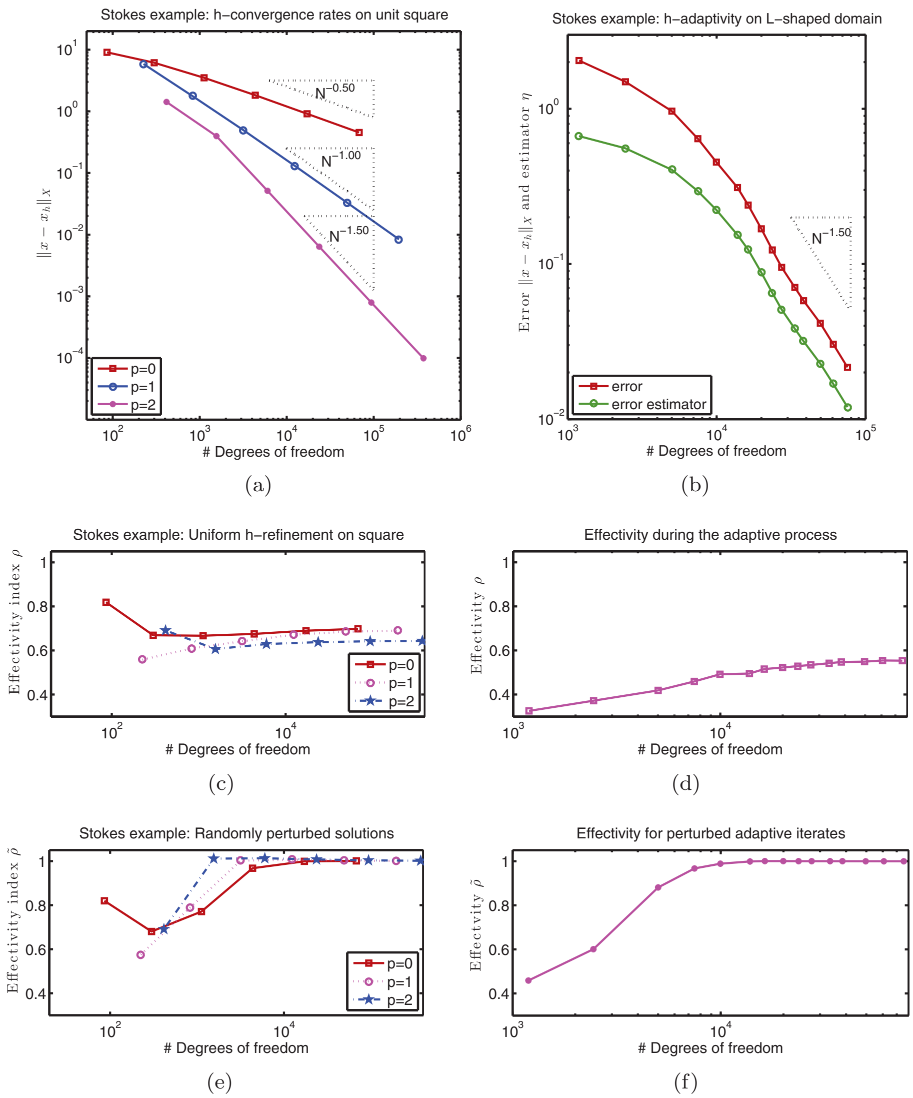

FIG. 3. *Results from the Stokes example for a smooth solution (left) and a singular solution (right).* (a) *Convergence rates for a smooth solution under uniform refinements of the unit square.* (b) *Error and estimator in an h-adaptive algorithm with $p = 2$ for the singular solution of Osborn.* (c) *Effectivity indices for the solutions of* (a). (d) *Effectivity indices for the solutions of* (b). (e) *Changes in effectivity when solutions of* (a) *are perturbed at each uniform refinement.* (f) *Changes in effectivity when solutions of* (b) *are perturbed at each adaptive refinement.*

REFERENCES

[1] M. Ainsworth, *A posteriori error estimation for discontinuous Galerkin finite element approximation*, SIAM J. Numer. Anal., 45 (2007), pp. 1777–1798.

[2] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley, Chichester, 2000.

[3] I. Babuška and T. Strouboulis, *The Finite Element Method and its Reliability*, Clarendon Press, Oxford, 2001.

[4] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations.* Lectures Math. ETH Zürich, Birkhäuser, Basel, 2003.
[5] P.B. Bochev and M.D. Gunzburger, *Least-Squares Finite Element Methods*, Appl. Math. Sci. 166, Springer, New York, 2009.
[6] A. Bonito and R.H. Nochetto, *Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method*, SIAM J. Numer. Anal., 48 (2010), pp. 734–771.
[7] J. Bramwell, L. Demkowicz, J. Gopalakrishnan, and W. Qiu, *A locking-free hp DPG method for linear elasticity with symmetric stresses*, Numer. Math., 122 (2012), pp. 671–707.
[8] Z. Cai and Y. Wang, *A multigrid method for the pseudostress formulation of Stokes problems*, SIAM J. Sci. Comput., 29 (2007), pp. 2078–2095.
[9] C. Carstensen, *A unifying theory of a posteriori finite element error control*, Numer. Math., 100 (2005), pp. 617–637.
[10] C. Carstensen, T. Gudi, and M. Jensen, *A unifying theory of a posteriori error control for discontinuous Galerkin FEM*, Numer. Math., 112 (2009), pp. 363–379.
[11] C. Carstensen, M. Eigel, R.H.W. Hoppe, and C. Löbhard, *A review of unified a posteriori finite element error control*, Numer. Math. Theory Methods Appl., 5 (2012), pp. 509–558.
[12] C. Carstensen, D. Kim, and E.-J. Park, *A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2501–2523.
[13] A. Cohen, W. Dahmen, and G. Welper, *Adaptivity and Variational Stabilization for Convection-Diffusion Equations*, Report 2011-323, Institut fuer Geometrie und Praktische Mathematik, Aachen, Germany.
[14] M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners*. Part I. *Linearized equations*, SIAM J. Math. Anal., 20 (1989), pp. 74–97.
[15] L. Demkowicz and J. Gopalakrishnan, *A class of discontinuous Petrov-Galerkin methods. Part* I: *The transport equation*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 1558–1572.
[16] L. Demkowicz and J. Gopalakrishnan, *A class of discontinuous Petrov-Galerkin methods. Part* II: *Optimal test functions*, Numer. Methods Partial Differential Equations, 27 (2011), pp. 70–105.
[17] L. Demkowicz and J. Gopalakrishnan, *Analysis of the DPG method for the Poisson equation*, SIAM J Numer. Anal., 49 (2011), pp. 1788–1809.
[18] L. Demkowicz and J. Gopalakrishnan, *A primal DPG method without a first-order reformulation*, Comput. Math. Appl., 66 (2013), pp. 1058–1064.
[19] L. Demkowicz, J. Gopalakrishnan, and A. Niemi, *A class of discontinuous Petrov-Galerkin methods. Part* III: *Adaptivity*, Appl. Numer. Math., 62 (2012), pp. 396–427.
[20] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, Cambridge 1995.
[21] J. Gopalakrishnan and J. Guzmán, *Symmetric nonconforming mixed finite elements for linear elasticity*, SIAM J Numer. Anal., 49 (2011), pp. 1504–1520.
[22] J. Gopalakrishnan and W. Qiu, *An analysis of the practical DPG method*, Math. Comp., 83 (2014), pp. 537–552,
[23] P. Grisvard, *Singularities in Boundary Value Problems*, Rech. Math. Appl. 22, Masson, Paris, 1992.
[24] P. Hansbo and M.G. Larson, *Energy norm a posteriori error estimates for discontinuous Galerkin approximations of the linear elasticity problem*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 3026–3030.
[25] P. Houston, D. Schötzau, and T.P. Wihler, *Energy norm a posteriori error estimation for mixed discontinuous Galerkin approximations of the Stokes problem*, J. Sci. Comput., 22/23 (2005), pp. 347–370.
[26] P. Houston, D. Schötzau, and T.P. Wihler, *An hp-adaptive mixed discontinuous Galerkin FEM for nearly incompressible linear elasticity*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3224–3246.
[27] O.A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.
[28] O.A. Karakashian and F. Pascal, *Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems*, SIAM J. Numer. Anal., 45 (2007), pp. 641–665.
[29] J. Osborn, *Regularity of the solutions of the Stokes problem in a polygonal domain*, in B. Hubbard, ed., Symposium on Numerical Solution Partial Differential Equations III, Academic, New York, pp. 393–411, 1975.

[30] N. Roberts, T. Bui-Thanh, and L. Demkowicz, *The DPG method for the Stokes problem*, Comput. Math. Appl., 67 (2014), pp. 966–995.

[31] T.P. Wihler, *Locking-free adaptive discontinuous Galerkin FEM for linear elasticity problems*, Math. Comp., 75 (2006), pp. 1087–1102.

[32] R. Verfürth, *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, New York, 1996.