

# Optimal BV Estimates for a Discontinuous Galerkin Method for Linear Elasticity

Adrian Lew, Patrizio Neff,  
Deborah Sulsky, and Michael Ortiz

## 1 Introduction

Discontinuous Galerkin (DG) finite-element methods for second- and fourth-order elliptic problems were introduced about three decades ago. These methods stem from the hybrid methods developed by Pian and his coworker [25]. At the time of their introduction, DG methods were generally called interior penalty methods, and were considered by Baker [4], Douglas Jr. [14], and Douglas Jr. and Dupont [15] for fourth-order problems, where  $C^1$  continuity was imposed on  $C^0$  elements. For second-order equations, Nitsche [21] appears to have introduced the ideas of imposing Dirichlet boundary conditions weakly and of adding stabilization terms to obtain optimal convergence rates. The same idea of penalizing jumps along interelement faces led to the interior penalty methods of Percell and Wheeler [24] and Wheeler [30]. Methods for a second-order, nonlinear, parabolic equation appeared in [1].

According to [3], interest in DG methods for solving elliptic problems waned because they were never proven to be more advantageous than traditional conforming elements. The difficulty in identifying optimal penalty parameters and efficient solvers may also have contributed to the lack of interest [3]. Recently, however, interest has been rekindled by developments in DG methods for convection-diffusion problems; see, for example, Cockburn and Shu [12, 13], Oden, Babuška, and Baumann [22], Castillo, Cockburn,

Received 2 February 2004.

Communicated by Thomas Yizhao Hou.

Perugia, and Schötzau [9], and Houston, Schwab, and Süli [18], where the scalar Poisson equation is analyzed. Bassi and Rebay [5] applied a similar technique for the solution of the Navier-Stokes equations. Brezzi, Manzini, Marini, Pietra, and Russo [7, 8] analyzed the method of Bassi and Rebay for stability and accuracy, as it applies to the scalar Poisson equation. Arnold, Brezzi, Cockburn, and Marini [2, 3] provided a common framework for all of these methods and showed the interconnections by casting them into the form of the local discontinuous Galerkin (LDG) method of Cockburn and Shu.

We are interested in a DG method for studying the mechanical behavior of solids. In this paper, we analyze the linear elasticity problem, with an eye toward a formulation for nonlinear elastic-plastic problems and cohesive elements [23]. There are several benefits of such an approach, including the potential for efficient hp-adaptivity, for example, using adaptive mesh refinement on meshes with hanging nodes, and the prospect of rigorously handling problems with discontinuous displacements as arise in the study of fracture. Rivière and Wheeler [26] formulate and analyze a method for linear elasticity based on a generalization of the nonsymmetric interior penalty Galerkin (NIPG) method presented in [22] for the diffusion equation. The resulting bilinear form is nonsymmetric. As an alternative, we follow the analysis of Brezzi, Manzini, Marini, Pietra, and Russo [7, 8] quite closely in our generalization from the scalar Poisson equation to the linear elasticity problem. In this case, the bilinear form is symmetric.

Error estimates for DG methods are usually obtained in terms of mesh-dependent norms. It is, a priori, not clear how to compare norms corresponding to meshes of different size. In this paper, we show that the traditional error estimates expressed in mesh-dependent norms can be used to derive error estimates in the mesh-independent BD and BV norms, eliminating the ambiguity.

Section 2 begins with a statement of the problem and its formulation using the DG approach. A new derivation of the equations is based on a discrete variational principle for elasticity which naturally extends to finite deformations. The variational approach leads to a formulation analogous to the one utilized in [5]. Stabilization terms of the form considered in [7, 8] are added to obtain a well-posed discrete problem. In Section 3, we show optimal convergence rates in a mesh-dependent norm similar to the one used by Brezzi et al. This mesh-dependent estimate is immediately strengthened to a mesh-independent BD estimate in Section 3.2.

The classical analysis of the equations of linear elasticity needs a global version of Korn's first inequality to insure coerciveness of the bilinear form. In contrast to the standard approach, in Section 3.3, we prove a generalization of Korn's second inequality on the element level, which allows us to obtain an improved mesh-dependent estimate. Finally, in Section 3.5, we show uniform convergence in the BV norm, an optimal

mesh-independent estimate. Since the discrete solutions are allowed to have jumps in displacement but the classical solution is smooth, gradients can at most converge in measure, and indeed they do.

## 2 Formulation

The linear elasticity problem is described by the following set of equations for a body  $B \subset \mathbb{R}^d$ , where  $d = 2, 3$ :

$$\begin{aligned} -\nabla \cdot (\mathbb{C} \cdot \nabla_s u) &= f \quad \text{in } B, \\ u &= \bar{u} \quad \text{on } \partial_D B, \\ (\mathbb{C} \cdot \nabla_s u) \cdot n &= \bar{T} \quad \text{on } \partial_N B. \end{aligned} \tag{2.1}$$

The body  $B$  is assumed to be a bounded, polyhedral domain. The function  $u : B \rightarrow \mathbb{R}^d$  is the displacement, and  $\mathbb{C}$  is the fourth-order elasticity tensor with major and minor symmetries. In order to avoid technical difficulties that do not provide any additional insight, we take  $\mathbb{C}$  to be constant. We also assume that  $\mathbb{C}$  is uniformly positive definite, that is,

$$\exists c > 0 : \gamma \cdot \mathbb{C} \cdot \gamma \geq c \gamma \cdot \gamma \tag{2.2}$$

for all  $\gamma$  in the space of  $d \times d$  symmetric tensors, which implies that  $\mathbb{C}$  is invertible on this space. The notation  $\nabla_s u$  denotes the symmetric gradient of the displacement,  $\nabla_s u = (1/2)(\nabla u + (\nabla u)^T)$ . The boundary of the domain,  $\partial B$ , is decomposed into two disjoint sets,  $\partial_D B$  and  $\partial_N B$ . The body is acted upon by body forces,  $f : B \rightarrow \mathbb{R}^d$ , and surface tractions,  $\bar{T} : \partial_N B \rightarrow \mathbb{R}^d$ . The displacement,  $\bar{u} : \partial_D B \rightarrow \mathbb{R}^d$ , is prescribed on the part of the boundary indicated by  $\partial_D B$ .

### 2.1 Stress-displacement formulation

The two-field, stress-displacement formulation of the linear elasticity problem is

$$\begin{aligned} \sigma - \mathbb{C} \cdot \nabla_s u &= 0 \quad \text{in } B, \\ -\nabla \cdot \sigma &= f \quad \text{in } B, \\ u &= \bar{u} \quad \text{on } \partial_D B, \\ \sigma \cdot n &= \bar{T} \quad \text{on } \partial_N B. \end{aligned} \tag{2.3}$$

The first equation is the constitutive equation that relates the stress tensor  $\sigma$  to the strain  $\varepsilon(\nabla u) = \nabla_s u$ . The second equation expresses force equilibrium, and the final two equations give the prescribed boundary conditions. The problem described by equation (2.3) has solutions  $(u, \sigma)$  with components in  $H^{m+1}(B)$  and  $H^m(B)$ , respectively, for  $m \geq 1$ , depending on the smoothness of the data and the domain. Nominally,  $f \in (L^2(B))^d$ .

The equations (2.3) are the Euler-Lagrange equations that result from taking free variations of the Hellinger-Reissner energy,  $I : (H^{m+1}(B))^d \times (H^m(B))^{d \times d} \rightarrow \mathbb{R}$ , where

$$\begin{aligned} I[u, \sigma] = & \int_B \left( \frac{1}{2} \sigma \cdot \mathbb{C}^{-1} \cdot \sigma - \sigma \cdot \nabla_s u + f \cdot u \right) \\ & + \int_{\partial_D B} n \cdot \sigma \cdot (u - \bar{u}) + \int_{\partial_N B} \bar{T} \cdot u. \end{aligned} \quad (2.4)$$

The discrete equations in the next section are derived using a discretization of this variational principle.

## 2.2 The discrete scheme

A subdivision,  $\mathcal{T}_h$  of  $B$ , is a finite number of sets  $E$ , such that  $\bar{B} = \cup_{E \in \mathcal{T}_h} E$ . A subdivision,  $\mathcal{T}_h$ , is called admissible in the sense of [10, page 38] if each  $E$  is closed and has nonempty interior, the interiors of the sets  $E$  of  $\mathcal{T}_h$  are pairwise disjoint, and the boundary,  $\partial E$ , of each  $E$  is Lipschitz continuous. We assume the family of admissible subdivisions  $(\mathcal{T}_h)$ , with  $h \downarrow 0$ , is quasi-uniform [6, page 106] so that

$$\max \{ \text{diam } E : E \in \mathcal{T}_h \} = h; \quad (2.5)$$

$$\exists \rho > 0 : \min \{ \text{diam } B_E : E \in \mathcal{T}_h \} \geq \rho h, \forall h > 0, \quad (2.6)$$

where  $B_E$  is the largest ball contained in  $E$ . Therefore, it follows that there exist positive constants  $c$  and  $C$  such that

$$ch^d \leq |E| \leq Ch^d \quad (2.7)$$

for every element  $E \in \mathcal{T}_h$  and every  $h > 0$ , where  $|E|$  is the measure of  $E$ . In addition, we require all finite elements within the family of subdivisions to be affine equivalent [6, page 80] to a finite number of polyhedral reference finite elements, each with a finite number of faces. Hence the reference elements possess Lipschitz boundaries, the measure of each face of an arbitrary element in  $(\mathcal{T}_h)$  is finite, and there exists an upper bound on the Lipschitz constant of the boundary for all elements in the family  $(\mathcal{T}_h)$ , independent of  $h$ .

Moreover, with (2.5), we infer that there exists a constant  $C > 0$  such that

$$|e|h \leq C|E|, \quad (2.8)$$

for all  $h > 0$ , and for any face  $e$  of any element  $E \in (\mathcal{T}_h)$ . Even though DG methods can potentially be used on meshes with hanging nodes, we consider, for simplicity, only conforming meshes, so that a face  $e$  of an element is either also a face of another element or part of  $\partial B$ . We note, however, that most of the theoretical development does not rely on this assumption.

Consider a given subdivision  $\mathcal{T}_h$  of  $B$ . Each element  $E \in \mathcal{T}_h$  has an orientable boundary,  $\partial E$ , with unit, outward normal denoted by  $n_E$ . Define the set of internal faces

$$\mathcal{E}_h^I = \{e \subset \partial E \setminus \partial B : E \in \mathcal{T}_h\}, \quad (2.9)$$

the set of Dirichlet faces

$$\mathcal{E}_h^D = \{e \subset \partial E \cap \partial_D B : E \in \mathcal{T}_h\}, \quad (2.10)$$

and the set of Neumann faces

$$\mathcal{E}_h^N = \{e \subset \partial E \cap \partial_N B : E \in \mathcal{T}_h\}. \quad (2.11)$$

The set of all faces is denoted by  $\mathcal{E}_h = \mathcal{E}_h^I \cup \mathcal{E}_h^D \cup \mathcal{E}_h^N$ . Corresponding to this set of faces, define the combined internal and external boundary to be

$$\Gamma = \cup_{e \in \mathcal{E}_h} e. \quad (2.12)$$

Let  $\tilde{V} = \Pi_{E \in \mathcal{T}_h} (H^1(E))^d$  be the space of functions on  $B$  whose restriction to each element  $E$  belongs to the Sobolev space  $(H^1(E))^d$ . Therefore, the traces of functions in  $\tilde{V}$  belong to  $T(\Gamma) = \Pi_{E \in \mathcal{T}_h} (L^2(\partial E))^d$ . Functions in  $T(\Gamma)$  are multivalued on  $\Gamma \setminus \partial B$  and single-valued on  $\partial B$ . The space  $(L^2(\Gamma))^d$  can be identified with the subspace of  $T(\Gamma)$  consisting of functions for which the possible multiple values agree on all internal faces. Similarly, let  $\tilde{W} = \Pi_{E \in \mathcal{T}_h} (H^1(E))^{d \times d}$  be the space of functions on  $B$  whose restriction to each element  $E$  belongs to the Sobolev space  $(H^1(E))^{d \times d}$ . A tensor  $\tau \in \tilde{W}$  has  $d^2$  components. The  $d^2$  traces, the components of  $\tau|_{\partial E}$ , are defined, and each belongs to  $L^2(\partial E)$ . In particular, the linear combination of traces  $\tau \cdot n_E$  is in  $T(\Gamma)$ .<sup>1</sup>

<sup>1</sup>The space of stresses  $\tilde{W}$  could be taken to be larger; however, this is unnecessary since we consider exact solutions  $(u, \sigma)$  in  $(H^2(B))^d \times (H^1(B))^{d \times d}$ .

Next, we introduce two finite-element spaces of scalar functions over an element  $E$ ,  $V_h^E$  and  $W_h^E$ , with  $V_h^E \subseteq W_h^E$ . These elemental spaces contain the polynomials and have minimal smoothness over the element,  $\mathcal{P}_k(E) \subseteq V_h^E, W_h^E \subset H^1(E)$ ,  $k \geq 1$ , where  $\mathcal{P}_k(E)$  denotes the space of polynomials of degree at most  $k$  on  $E$ . The finite-element spaces for the displacements,  $V_h$ , and displacement gradients,  $W_h$ , are constructed so that each component is in  $V_h^E$  or  $W_h^E$  on the element  $E$ ,  $V_h = \Pi_{E \in \mathcal{T}_h}(V_h^E)^d$  and  $W_h = \Pi_{E \in \mathcal{T}_h}(W_h^E)^{d \times d}$ . Consequently, we have  $V_h \subset \tilde{V}$ . We also assume that gradients of the displacement are in the space of displacement gradients,  $\nabla[(V_h^E)^d] \subseteq (W_h^E)^{d \times d}$ . Furthermore, we require the elemental finite-element spaces to coincide over common faces. More precisely, let  $e \in \mathcal{E}_h^I$  be the face common to two elements,  $E^+$  and  $E^-$ , then  $\{\phi|_e : \phi \in V_h^{E^+}\} = \{\phi|_e : \phi \in V_h^{E^-}\}$  and  $\{\phi|_e : \phi \in W_h^{E^+}\} = \{\phi|_e : \phi \in W_h^{E^-}\}$ . This requirement insures that the trace of a function in  $V_h^{E^+}(W_h^{E^+})$  is also the trace of a function in  $V_h^{E^-}(W_h^{E^-})$ , on  $e$ . Lastly, we denote with  $W_h^s$  the space of symmetric tensors in  $W_h$ .

We assume that the discrete spaces,  $V_h$  and  $W_h$ , are finite dimensional. Observe that the functions in both discrete spaces can be discontinuous across element boundaries. The conditions specified here are satisfied by many standard finite-element spaces, such as those constructed from Lagrange simplices of various degrees and those constructed with bilinear quadrilaterals or trilinear bricks.

**Remark 2.1.** Most of the proofs in this article immediately generalize to the case of isoparametric elements, though some adjustment of the assumptions on the finite-element spaces might be required. In particular, the special treatment of Korn's inequality also applies to isoparametric elements.

We wish to formulate a discretized version of (2.4) subordinate to the subdivision. To this end, we define the average operator,  $\{\cdot\} : T(\Gamma) \rightarrow (L^2(\Gamma))^d$ , and the jump operator,  $[[\cdot]] : T(\Gamma) \rightarrow (L^2(\Gamma))^d$ . Each face,  $e \in \mathcal{E}_h^I$ , is shared by two elements,  $E^+$  and  $E^-$ ; let  $v^\pm = v|_{E^\pm}$  for  $v \in \tilde{V}$ . Define the average, for  $e \in \mathcal{E}_h^I$ , by

$$\{v\} = \frac{1}{2}(v^-|_e + v^+|_e) \quad (2.13)$$

and the jump by

$$[[v]] = v^-|_e - v^+|_e. \quad (2.14)$$

For  $e \in \mathcal{E}_h^D$ , put

$$\{v\} = v, \quad [[v]] = v; \quad (2.15)$$

and for  $e \in \mathcal{E}_h^N$ , assign

$$\{v\} = v, \quad \llbracket v \rrbracket = 0. \quad (2.16)$$

In the sequel, we choose an orientation,  $n$ , for each face  $e \in \mathcal{E}_h^I$ , as the unit normal pointing toward  $E^+$ . For  $e \subset \partial B$ ,  $n$  is the unit outward normal to  $\partial B$ . For  $\sigma \in \tilde{W}$ , let  $\sigma^\pm = \sigma|_{E^\pm}$ . On  $e \in \mathcal{E}_h^I$ , the average of the vector  $\sigma \cdot n$  means

$$\{\sigma \cdot n\} = \frac{1}{2}(\sigma^+|_e + \sigma^-|_e) \cdot n, \quad (2.17)$$

with  $n$  given uniquely on the face. The definition of  $\{\sigma \cdot n\}$  on boundary faces,  $e \in \mathcal{E}_h^D \cup \mathcal{E}_h^N$ , is clear.

Now, specialize (2.4) to each individual element as follows:

$$\begin{aligned} I_E = & \int_E \left( \frac{1}{2} \sigma \cdot \mathbb{C}^{-1} \cdot \sigma - \sigma \cdot \nabla_s u + f \cdot u \right) + \int_{\partial E \setminus \partial B} \frac{1}{2} n_E \cdot \sigma \cdot (u - u^{\text{ext}}) \\ & + \int_{\partial E \cap \partial_D B} n \cdot \sigma \cdot (u - \bar{u}) + \int_{\partial E \cap \partial_N B} \bar{T} \cdot u, \end{aligned} \quad (2.18)$$

where  $u^{\text{ext}}$  is the trace of  $u$  on the elements adjacent to  $\partial E \setminus \partial B$ . The  $1/2$  factor in the second term accounts for the fact that for a given face, two adjacent elements contribute to the potential energy. A global discrete functional,  $I_h : V_h \times W_h^s \rightarrow \mathbb{R}$ , is defined simply by summing over all elemental contributions:

$$I_h = \sum_{E \in \mathcal{T}_h} I_E. \quad (2.19)$$

The corresponding Euler-Lagrange equations that result from taking free variations of  $I_h$  are

$$\begin{aligned} \sum_{E \in \mathcal{T}_h} \int_E (\delta \sigma \cdot \mathbb{C}^{-1} \cdot \sigma - \delta \sigma \cdot \nabla_s u) + \int_\Gamma \{n \cdot \delta \sigma\} \cdot \llbracket u \rrbracket - \int_{\partial_D B} n \cdot \delta \sigma \cdot \bar{u} &= 0, \\ \sum_{E \in \mathcal{T}_h} \int_E (-\sigma \cdot \nabla_s \delta u + f \cdot \delta u) + \int_\Gamma \{n \cdot \sigma\} \cdot \llbracket \delta u \rrbracket + \int_{\partial_N B} \bar{T} \cdot \delta u &= 0. \end{aligned} \quad (2.20)$$

Thus, we obtain the general problem which is to find  $u_h \in V_h$  and  $\sigma_h \in W_h^s$  such that

$$\begin{aligned} \sum_{E \in \mathcal{T}_h} \int_E (\gamma_h \cdot \mathbb{C}^{-1} \cdot \sigma_h - \gamma_h \cdot \nabla_s u_h) + \int_{\Gamma} \{n \cdot \gamma_h\} \cdot \llbracket u_h \rrbracket \\ = \int_{\partial_D B} n \cdot \gamma_h \cdot \bar{u} \quad \forall \gamma_h \in W_h^s; \end{aligned} \quad (2.21)$$

$$\begin{aligned} \sum_{E \in \mathcal{T}_h} \int_E \sigma_h \cdot \nabla_s v_h - \int_{\Gamma} \{n \cdot \sigma_h\} \cdot \llbracket v_h \rrbracket \\ = \int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h \quad \forall v_h \in V_h. \end{aligned} \quad (2.22)$$

Equations (2.21) and (2.22) constitute the flux form of the discrete problem. Next, we define the lifting operator  $R_{\bar{u}} : (L^2(\Gamma))^d \rightarrow W_h^s$  by

$$\int_B R_{\bar{u}}(v) \cdot \gamma = - \int_{\Gamma} \{n \cdot \gamma\} \cdot v + \int_{\partial_D B} n \cdot \gamma \cdot \bar{u} \quad \forall \gamma \in W_h^s. \quad (2.23)$$

This operator will now be used to derive the primal form [3] of the discretization, where a single equation is obtained by eliminating  $\sigma_h$  between (2.21) and (2.22). In terms of (2.23), equation (2.21) is the same as

$$\sum_{E \in \mathcal{T}_h} \int_E (\gamma_h \cdot \mathbb{C}^{-1} \cdot \sigma_h - \gamma_h \cdot \nabla_s u_h - R_{\bar{u}}(\llbracket u_h \rrbracket) \cdot \gamma_h) = 0 \quad \forall \gamma_h \in W_h^s. \quad (2.24)$$

Since we require the elemental finite-element spaces to satisfy  $\nabla[(V_h^E)^d] \subseteq (W_h^E)^{d \times d}$ , this equation allows us to identify

$$\sigma_h = \sigma_h(u_h) = \mathbb{C} \cdot \nabla_s u_h + \mathbb{C} \cdot R_{\bar{u}}(\llbracket u_h \rrbracket) \quad \text{in } W_h^s. \quad (2.25)$$

This constitutive equation for the discrete stress can be viewed as a stress-strain relation where the strain involves the usual dependence on the displacement gradient, plus a linear contribution that arises from jumps in displacement.

Next, take  $\gamma_h = \mathbb{C} \cdot \nabla_s v_h$  in equation (2.21) to get

$$\begin{aligned} \sum_{E \in \mathcal{T}_h} \int_E (\nabla_s v_h \cdot \sigma_h - \nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u_h) + \int_{\Gamma} \{n \cdot \mathbb{C} \cdot \nabla_s v_h\} \cdot \llbracket u_h \rrbracket \\ = \int_{\partial_D B} n \cdot (\mathbb{C} \cdot \nabla_s v_h) \cdot \bar{u}. \end{aligned} \quad (2.26)$$



Finally, substitute equation (2.22) to obtain

$$\begin{aligned} & \sum_{E \in \mathcal{T}_h} \int_E \nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u_h - \int_{\Gamma} (\{n \cdot \mathbb{C} \cdot \nabla_s v_h\} \cdot \llbracket u_h \rrbracket + \{n \cdot \sigma_h\} \cdot \llbracket v_h \rrbracket) \\ &= \int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h - \int_{\partial_D B} n \cdot (\mathbb{C} \cdot \nabla_s v_h) \cdot \bar{u}. \end{aligned} \quad (2.27)$$

If  $(u_h, \sigma_h) \in V_h \times W_h^s$  solves (2.21) and (2.22), then  $u_h$  solves (2.27), with  $\sigma_h = \sigma_h(u_h)$  given by (2.25). Equation (2.27) is called the primal formulation.

Recall the definition of  $R_{\bar{u}}$  in (2.23) and introduce the notation  $R = R_0$ . Using (2.23) and (2.25), the primal form (2.27) can also be written as

$$\begin{aligned} & \sum_{E \in \mathcal{T}_h} \int_E (\nabla_s v_h + R(\llbracket v_h \rrbracket)) \cdot \mathbb{C} \cdot (\nabla_s u_h + R(\llbracket u_h \rrbracket)) \\ &= \int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h - \int_{\partial_D B} n \cdot (\mathbb{C} \cdot (\nabla_s v_h + R(\llbracket v_h \rrbracket))) \cdot \bar{u}. \end{aligned} \quad (2.28)$$

We remark that our physically based derivation of this equation, obtained by discretizing the variational principle, produces an analogous discretization to that used by Bassi and Rebay in [5, 8]. Brezzi, Manzini, Marini, Pietra, and Russo [7, 8] propose a stabilizing term for the scalar case which naturally extends to linear elasticity. The stabilization is given in terms of  $r_{e,\bar{u}} : (L^2(\Gamma))^d \rightarrow W_h^s$ . Define  $r_{e,\bar{u}}$  for  $e \in \mathcal{E}_h^I$ ,

$$\int_B r_{e,\bar{u}}(v) \cdot \gamma = - \int_e \{n \cdot \gamma\} \cdot v \quad \forall \gamma \in W_h^s, \quad (2.29)$$

while for  $e \in \mathcal{E}_h^D$ ,

$$\int_B r_{e,\bar{u}}(v) \cdot \gamma = - \int_e \{n \cdot \gamma\} \cdot v + \int_e n \cdot \gamma \cdot \bar{u} \quad \forall \gamma \in W_h^s, \quad (2.30)$$

and for  $e \in \mathcal{E}_h^N$ ,  $r_{e,\bar{u}} = 0$ . As before, set  $r_e = r_{e,0}$ . Note that  $r_{e,\bar{u}}(v)$  vanishes outside the union of elements containing  $e$ , and that for any element  $E \in \mathcal{T}_h$ ,

$$R_{\bar{u}}(v) = \sum_{e \subset \partial E} r_{e,\bar{u}}(v) \quad (2.31)$$

on  $E$ . The stabilizing term is  $\beta \sum_{e \subset \partial E} \int_B r_{e,\bar{u}}(\llbracket u_h \rrbracket) \cdot \mathbb{C} \cdot r_e(\llbracket v_h \rrbracket)$ , with  $\beta > 0$  the stabilization parameter. The resulting primal form with the stabilizing term is

$$\begin{aligned} & \sum_{E \in \mathcal{T}_h} \int_E (\nabla_s v_h + R(\llbracket v_h \rrbracket)) \cdot \mathbb{C} \cdot (\nabla_s u_h + R(\llbracket u_h \rrbracket)) + \beta \sum_{e \in \mathcal{E}_h} \int_B r_e(\llbracket u_h \rrbracket) \cdot \mathbb{C} \cdot r_e(\llbracket v_h \rrbracket) \\ &= \int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h - \int_{\partial_D B} n \cdot (\mathbb{C} \cdot (\nabla_s v_h + R(\llbracket v_h \rrbracket)) + \beta r_e(\llbracket v_h \rrbracket))) \cdot \bar{u}. \end{aligned} \quad (2.32)$$

The form (2.32), which derives directly from the variational principle, is stable for any  $\beta > 0$ . In Section 3, we analyze in detail a modification proposed by Brezzi, Manzini, Marini, Pietra, and Russo [7, 8] that omits the quadratic term in  $R$ , making the method stable for  $\beta > N_e$ , where  $N_e$  is the maximum number of faces in an element of the subdivision. The advantage of dropping this quadratic term is that the sparsity of the stiffness matrix is increased.

The analysis of the proposed method relies on elliptic regularity, so we restrict it to Dirichlet boundary conditions on the entire boundary,  $\partial B$ . Thus,  $\partial_N B = \emptyset$ ,  $\mathcal{E}_h^N = \emptyset$ , and, without loss of generality,  $\bar{u} = 0$  on  $\partial B$ . Accordingly, the complete discrete problem statement, with these modifications, is to find  $u_h \in V_h$  such that

$$a_h(u_h, v_h) = \int_B f \cdot v_h \quad \forall v_h \in V_h, \quad (2.33)$$

where the bilinear form  $a_h$  is given by

$$\begin{aligned} a_h(u_h, v_h) = & \sum_{E \in \mathcal{T}_h} \int_E (\nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u_h + \nabla_s v_h \cdot \mathbb{C} \cdot R(\llbracket u_h \rrbracket) + R(\llbracket v_h \rrbracket) \cdot \mathbb{C} \cdot \nabla_s u_h) \\ & + \beta \sum_{e \in \mathcal{E}_h} \int_B r_e(\llbracket u_h \rrbracket) \cdot \mathbb{C} \cdot r_e(\llbracket v_h \rrbracket). \end{aligned} \quad (2.34)$$

Remark 2.2. The bilinear form (2.34) can be written alternatively using (2.23) as

$$\begin{aligned} a_h(u_h, v_h) = & \sum_{E \in \mathcal{T}_h} \int_E (\nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u_h) \\ & - \int_\Gamma (\{n \cdot \mathbb{C} \cdot \nabla_s v_h\} \cdot \llbracket u_h \rrbracket + \{n \cdot \mathbb{C} \cdot \nabla_s u_h\} \cdot \llbracket v_h \rrbracket) \\ & + \beta \sum_{e \in \mathcal{E}_h} \int_B r_e(\llbracket u_h \rrbracket) \cdot \mathbb{C} \cdot r_e(\llbracket v_h \rrbracket). \end{aligned} \quad (2.35)$$

These two forms are equivalent for  $u_h, v_h \in V_h$ .

### 2.3 Notation

In Section 3, a convergence proof will be given for  $d = 2$  and  $3$ , simultaneously. In the proofs, the letter  $C$  indicates a generic constant whose value can change in each occurrence. We also employ the standard notation  $\|\cdot\|_{p,\Omega}$  to denote the usual norm on  $H^p(\Omega)$ ,

and  $|\cdot|_{p,\Omega}$  to denote the  $H^p(\Omega)$  seminorm, whereas  $\|\cdot\|$  denotes the Euclidean norm for vectors or tensors. When other standard norms are used, they will be indicated explicitly with a subscript; for example,  $\|\cdot\|_{L^1(\Omega)}$  indicates the  $L^1(\Omega)$ -norm.

## 2.4 Summary of the theoretical results

The convergence proof utilizes two relevant mesh-dependent norms on  $\widehat{V} = (H_0^1(B))^d + V_h$  given by

$$|||v|||_s^2 = \sum_{E \in \mathcal{T}_h} \|\nabla_s v\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket v \rrbracket)\|_{0,B}^2, \quad v \in \widehat{V}, \quad (2.36)$$

$$|||v|||^2 = \sum_{E \in \mathcal{T}_h} \|\nabla v\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket v \rrbracket)\|_{0,B}^2, \quad v \in \widehat{V}. \quad (2.37)$$

**Proposition 3.4** establishes that  $|||\cdot|||_s$  is a norm on  $\widehat{V}$ . Also note that

$$|||v|||_s^2 \leq |||v|||^2, \quad v \in \widehat{V}, \quad (2.38)$$

which shows that  $|||\cdot|||$  is also a norm on  $\widehat{V}$ . Although one might expect the second term in the definitions of norms (2.36) and (2.37) to act as an  $L^2$ -like contribution, we can only assert that these are seminorms on  $\widehat{V}$ . In the case of the scalar Poisson equation [2, 3, 7, 8, 9, 18, 22], there is no need to distinguish between the norms (2.36) and (2.37). Following the ideas in [7, 8], it is straightforward to obtain boundedness and coercivity of the bilinear form  $a_h$  with respect to the mesh-dependent norm,  $|||\cdot|||_s$  (Proposition 3.5), which leads to convergence of the discrete solutions in the  $|||\cdot|||_s$ -norm and in  $L^2(B)$  (Theorems 3.14 and 3.15). The convergence in the  $|||\cdot|||_s$ -norm is sufficient for a mesh-independent BD estimate (Theorem 3.19); however, the  $|||\cdot|||_s$  norm does not provide control over the antisymmetric part of the displacement gradient. If the displacements are in  $H_0^1(B)$ , the equivalence of the two norms,  $|||\cdot|||_s$  and  $|||\cdot|||$ , relies on Korn's first inequality; for nonconforming elements, Korn's inequality may not be valid [17].

order to obtain convergence in the norm,  $|||\cdot|||$ , Theorem 3.23, we prove a generalized version of Korn's second inequality for the subdivision, Corollary 3.22. The proof of this inequality relies on observations about how Korn's inequality for an element behaves under distortion (Theorem 3.20) and scaling (Theorem 3.21). Finally, Theorem 3.26 shows that the mesh-dependent norm,  $|||\cdot|||$ , estimates the BV norm, and as a consequence, Corollary 3.27, we obtain convergence in BV, an optimal mesh-independent result.

### 3 Theoretical results

#### 3.1 Convergence in the mesh-dependent symmetric norm

In this section, we obtain the convergence of the discretized solutions in the mesh-dependent norm  $||| \cdot |||_s$ . Our analysis follows the outline in [7, 8] for the two-dimensional Poisson equation, but the details of most proofs differ. The first three lemmas characterize properties of the jumps. In the subsequent proposition, our analysis starts by establishing that  $||| \cdot |||_s$  is in fact a norm on  $\hat{V}$ .

**Lemma 3.1** (extension of traces). *Let  $e$  be a face of an element  $E \in (\mathcal{T}_h)$ . For any  $\phi$  in the trace space,  $T(e) = \{\phi \in (L^2(e))^{d \times d} : \phi = \gamma|_e, \gamma \in (W_h^E)^{d \times d}\}$ , there exists  $P_e(\phi) \in (W_h^E)^{d \times d}$  such that  $P_e(\phi)|_e = \phi$ . Moreover, for all  $\phi \in T(e)$ ,*

$$\exists C > 0 : \|P_e(\phi)\|_{0,E} \leq Ch^{1/2} \|\phi\|_{0,e} \quad (3.1)$$

for all  $h > 0$  and for all  $E \in \mathcal{T}_h$ . □

*Proof.* First examine a reference element. Let  $\hat{e} \subset \partial \hat{E}$  be a face of one of the reference elements,  $\hat{E}$ , and let  $\phi \in T(\hat{e})$ . There exists  $C > 0$  such that

$$\sup_{\phi \in T(\hat{e}), \|\phi\|_{0,\hat{e}}=1} \inf_{\gamma_h \in (W_h^{\hat{E}})^{d \times d}, \gamma_h|_{\hat{e}}=\phi} \|\gamma_h\|_{0,\hat{E}}^2 < C. \quad (3.2)$$

Since  $\gamma_h \in (W_h^{\hat{E}})^{d \times d}$  is a linear combination of basis functions on  $\hat{E}$ ,  $\|\gamma_h\|_{0,\hat{E}}$  is a quadratic form in a finite-dimensional space. Therefore, there is a minimizer,  $P_{\hat{e}}(\phi)$ , of  $\|\gamma_h\|_{0,\hat{E}}^2$  subject to the linear constraint  $\gamma_h|_{\hat{e}} = \phi \in T(\hat{e})$ , which depends continuously on  $\phi$ . Thus,  $P_{\hat{e}}(\phi)$  is bounded on the compact set  $\|\phi\|_{0,\hat{e}} = 1$ , and (3.2) follows.

Next, note that  $P_{\hat{e}}(\lambda\phi) = \lambda P_{\hat{e}}(\phi)$  for  $\lambda \in \mathbb{R}$ , which implies

$$\|P_{\hat{e}}(\phi)\|_{0,\hat{E}}^2 \leq C \|\phi\|_{0,\hat{e}}^2 \quad (3.3)$$

for all  $\phi \in T(\hat{e})$ . Since the number of reference elements is finite, as is the number of faces per element, we can choose  $C$  in (3.3) independent of the reference element and the face.

Now, let  $E$  be any element in the family of subdivisions  $(\mathcal{T}_h)$ , and let  $e$  be any one of its faces. Let  $F$  be the affine transformation such that  $E = F(\hat{E})$  for one of the reference elements  $\hat{E}$ , and let  $\hat{e}$  be the corresponding face in the reference element,  $e = F(\hat{e})$ . Given

$\phi \in T(e)$ , the definition of affine equivalence implies  $\hat{\phi} = \phi \circ F \in T(\hat{e})$ . Define  $P_e(\phi) = P_{\hat{e}}(\hat{\phi}) \circ F^{-1} \in (W_h^E)^{d \times d}$ , and note  $P_e(\phi)|_e = \phi$ . Then, use (3.3) and  $\|F\| \leq h/\hat{\rho}$  (see, e.g., [10, page 120]), where  $\hat{\rho}$  is the diameter of the largest ball contained in  $\hat{E}$ , to obtain

$$\begin{aligned}
 \int_E |P_e(\phi)|^2 &= |\det F| \int_{\hat{E}} |P_{\hat{e}}(\hat{\phi})|^2 \\
 &\leq C |\det F| \int_{\hat{e}} |\hat{\phi}|^2 \\
 &\leq C |\det F| \int_e |\phi|^2 |\det F^{-1}| \|F\| \\
 &\leq C \|F\| \int_e |\phi|^2 \\
 &\leq C \frac{h}{\hat{\rho}} \int_e |\phi|^2.
 \end{aligned} \tag{3.4}$$

The lemma follows. ■

**Lemma 3.2** (trace inequality for  $r_e$ ). There exists a constant  $C > 0$ , independent of the face  $e \in \mathcal{E}_h$  and of  $h$ , such that

$$\|r_e(v)\|_{0,e} \leq Ch^{-1/2} \|r_e(v)\|_{0,E} \tag{3.5}$$

for all  $v \in (L^2(e))^d$ . □

**Proof.** The inequality (3.5) is actually a statement about tensors  $\gamma \in (W_h^E)^{d \times d}$ , where  $\gamma = r_e(v)$ . The proof follows a scaling argument. Let  $\hat{e} \subset \partial \hat{E}$  be a face of one of the reference elements,  $\hat{E}$ . Then, there exists a constant  $C > 0$  such that

$$\|\hat{\gamma}\|_{0,\hat{e}} \leq C \|\hat{\gamma}\|_{0,\hat{E}} \tag{3.6}$$

for all  $\hat{\gamma} \in (W_h^{\hat{E}})^{d \times d}$ . Inequality (3.6) is a direct consequence of the continuity of the trace in  $W_h^{\hat{E}} \subset H^1(\hat{E})$  (see, e.g., [6, page 37]) and the fact that in a finite-dimensional space, all norms are equivalent. Since there are a finite number of reference elements, each with a finite number of faces, the constant  $C$  can be chosen independent of the reference element and of its face.

Now, consider  $\gamma \in (W_h^E)^{d \times d}$ , where  $E$  is an element affine equivalent to  $\hat{E}$ . Then, there exists an affine mapping  $F$  such that  $E = F(\hat{E})$ , and  $\hat{\gamma} \in (W_h^{\hat{E}})^{d \times d}$  such that  $\gamma = \hat{\gamma} \circ F^{-1}$ . Note that

$$\begin{aligned}
 \|\gamma\|_{0,E}^2 &= \int_E \gamma \cdot \gamma = |\det F| \int_{\hat{E}} \hat{\gamma} \cdot \hat{\gamma} = |\det F| \|\hat{\gamma}\|_{0,\hat{E}}^2, \\
 \|\gamma\|_{0,e}^2 &= \int_e \gamma \cdot \gamma = \|F^{-1} \hat{n}\| |\det F| \int_{\hat{e}} \hat{\gamma} \cdot \hat{\gamma} \leq \|F^{-1}\| |\det F| \|\hat{\gamma}\|_{0,\hat{e}}^2,
 \end{aligned} \tag{3.7}$$

where  $\hat{n}$  is the unit outward normal to  $\hat{e}$ . Therefore, (3.6) and (3.7) combine to yield

$$\|\gamma\|_{0,e} \leq C \|F^{-1}\|^{1/2} \|\gamma\|_{0,E} \leq \frac{C \hat{h}^{1/2}}{\rho^{1/2}} h^{-1/2} \|\gamma\|_{0,E}. \quad (3.8)$$

The last part of the bound uses the fact that  $\|F^{-1}\| \leq \hat{h}/(\text{diam } B_E) \leq \hat{h}/(\rho h)$  (see, e.g., [10, page 120]). ■

**Lemma 3.3** (jump bound). There exist two positive constants  $C_1$  and  $C_2$ , independent of the face  $e \in \mathcal{E}_h$  and of  $h$ , such that

$$\|[\![v_h]\!]\|_{0,e} \leq C_1 h^{1/2} \|r_e([\![v_h]\!])\|_{0,B} \quad \forall v_h \in V_h; \quad (3.9)$$

$$\|r_e([\![v_h]\!])\|_{0,B} \leq C_2 h^{-1/2} \|[\![v_h]\!]\|_{0,e} \quad \forall v_h \in V_h. \quad (3.10)$$

□

**Proof.** Let  $e \subset E$  be a face of element  $E$ . Given  $[\![v_h]\!] \in (L^2(e))^d$ , let  $\gamma_h^e \in (L^2(e))^{d \times d}$  be such that  $\gamma_h^e \cdot n = [\![v_h]\!]$ . Note that it is possible to choose  $\gamma_h^e$  so that  $\|\gamma_h^e\| \leq C \|[\![v_h]\!]\|$ . For the tensor  $\gamma_h^e$  defined only on  $e$ , construct an extension to the element,  $\gamma_h|_E = P_e(\gamma_h^e)$ , as in Lemma 3.1. Take  $\gamma_h \in W_h^s$  to be  $\gamma_h|_E = P_e(\gamma_h^e)$  on  $E$ ,  $\gamma_h = 0$  elsewhere, and  $v = [\![v_h]\!]$  in equation (2.29) to get

$$\begin{aligned} \frac{1}{2} \|[\![v_h]\!]\|_{0,e}^2 &= \frac{1}{2} \int_e [\![v_h]\!] \cdot [\![v_h]\!] \\ &\leq \int_B |r_e([\![v_h]\!]) \cdot P_e(\gamma_h^e)| \\ &\leq \|r_e([\![v_h]\!])\|_{0,B} \|P_e(\gamma_h^e)\|_{0,E} \\ &\leq C h^{1/2} \|r_e([\![v_h]\!])\|_{0,B} \|[\![v_h]\!]\|_{0,e}. \end{aligned} \quad (3.11)$$

In the nontrivial case in which  $\|[\![v_h]\!]\|_{0,e} \neq 0$ , inequality (3.9) follows from (3.11) by dividing through by  $\|[\![v_h]\!]\|_{0,e}$ .

To prove (3.10), take  $\gamma = r_e([\![v_h]\!])$  and  $v = [\![v_h]\!]$  in equation (2.29) to get

$$\begin{aligned} \|r_e([\![v_h]\!])\|_{0,B}^2 &= \left| \int_e \{n \cdot r_e([\![v_h]\!])\} \cdot [\![v_h]\!] \right| \\ &\leq \|[\![v_h]\!]\|_{0,e} \|\{r_e([\![v_h]\!])\}\|_{0,e} \\ &\leq C_2 h^{-1/2} \|[\![v_h]\!]\|_{0,e} \|r_e([\![v_h]\!])\|_{0,B}. \end{aligned} \quad (3.12)$$

We have used the linearity of  $r_e$  and (3.5) in the last step. The result, (3.10), follows. ■

**Proposition 3.4** (symmetric norm). Let  $v_h \in \widehat{V} = (H_0^1(B))^d + V_h$ . Then  $||| \cdot |||_s : \widehat{V} \rightarrow \mathbb{R}$  as defined in (2.36) is a norm on  $\widehat{V}$ .  $\square$

*Proof.* It is immediate that  $|||\lambda v|||_s = |\lambda| |||v|||_s$  for all  $\lambda \in \mathbb{R}$ , and that the triangle inequality holds since  $r_e$  is linear. We show that  $|||v|||_s = 0$  implies  $v = 0$  in  $\widehat{V}$ . Notice that  $|||v|||_s = 0$  if and only if  $\|\nabla_s v\|_{0,E} = 0$  for all  $E \in \mathcal{T}_h$  and  $\|r_e(\llbracket v \rrbracket)\|_{0,B} = 0$  for all  $e \in \mathcal{E}_h$ . Let  $v = v_1 + v_2 \in \widehat{V}$ , with  $v_1 \in (H_0^1(B))^d$  and  $v_2 \in V_h$ . By Lemma 3.3, we have that  $\|\llbracket v_2 \rrbracket\|_{0,e} \leq Ch^{1/2} \|r_e(\llbracket v_2 \rrbracket)\|_{0,B}$ . Therefore,  $\|\llbracket v_2 \rrbracket\|_{0,e} = 0$ . Since also  $\|\llbracket v_1 \rrbracket\|_{0,e} = 0$ , we have  $\|\llbracket v \rrbracket\|_{0,e} = 0$ . So  $v \in (H_0^1(B))^d$  by [27, Theorem 1.3]. Korn's first inequality for homogeneous boundary data applied to  $v \in (H_0^1(B))^d$  then shows that  $v = 0$ .  $\blacksquare$

Next, we show that the bilinear form (2.34) is continuous and coercive with respect to the norm,  $||| \cdot |||_s$ . The proofs follow [7, 8] almost exactly.

**Proposition 3.5** (continuity and coercivity of the bilinear form). Let  $N_e$  be a bound on the number of faces in an element. Then, there exists a constant  $M > 0$ , independent of  $h$ , such that

$$(i) \quad a_h(u_h, v_h) \leq M |||u_h|||_s |||v_h|||_s \text{ for all } u_h, v_h \in \widehat{V}.$$

Moreover, for  $\beta > N_e$ , there exists a constant  $\mu > 0$ , independent of  $h$ , such that

$$(ii) \quad a_h(u_h, u_h) \geq \mu |||u_h|||_s^2 \text{ for all } u_h \in \widehat{V}. \quad \square$$

*Proof.* We first prove the following inequality, a consequence of equation (2.31):

$$\|R(\llbracket v_h \rrbracket)\|_{0,E}^2 \leq N_e \sum_{e \subset \partial E} \|r_e(\llbracket v_h \rrbracket)\|_{0,E}^2. \quad (3.13)$$

We have

$$\begin{aligned} \|R(\llbracket v_h \rrbracket)\|_{0,E}^2 &= \int_E \left( \sum_{e \subset \partial E} r_e(\llbracket v_h \rrbracket) \right) \left( \sum_{e' \subset \partial E} r_{e'}(\llbracket v_h \rrbracket) \right) \\ &\leq \int_E \sum_{e' \subset \partial E} \sum_{e \subset \partial E} \|r_e(\llbracket v_h \rrbracket)\| \|r_{e'}(\llbracket v_h \rrbracket)\| \\ &\leq \sum_{e' \subset \partial E} \sum_{e \subset \partial E} \frac{1}{2} \left( \|r_e(\llbracket v_h \rrbracket)\|_{0,E}^2 + \|r_{e'}(\llbracket v_h \rrbracket)\|_{0,E}^2 \right) \\ &\leq N_e \sum_{e \subset \partial E} \|r_e(\llbracket v_h \rrbracket)\|_{0,E}^2. \end{aligned} \quad (3.14)$$

Next, the continuity of the bilinear form (2.34) follows from estimating each term.

$$\begin{aligned}
\left| \int_E \nabla_s \mathbf{u}_h \cdot \mathbb{C} \cdot \nabla_s \mathbf{v}_h \right| &\leq \|\mathbb{C}\| \|\nabla_s \mathbf{u}_h\|_{0,E} \|\nabla_s \mathbf{v}_h\|_{0,E}, \\
\left| \int_E \nabla_s \mathbf{u}_h \cdot \mathbb{C} \cdot \mathbf{R}(\llbracket \mathbf{v}_h \rrbracket) \right| &\leq \|\mathbb{C}\| \|\nabla_s \mathbf{u}_h\|_{0,E} \|\mathbf{R}(\llbracket \mathbf{v}_h \rrbracket)\|_{0,E} \\
&\leq \|\mathbb{C}\| \|\nabla_s \mathbf{u}_h\|_{0,E} \left[ N_e \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket \mathbf{v}_h \rrbracket)\|_{0,E}^2 \right]^{1/2}, \\
\left| \sum_{e \subset \partial E} \int_E \mathbf{r}_e(\llbracket \mathbf{u}_h \rrbracket) \cdot \mathbb{C} \cdot \mathbf{r}_e(\llbracket \mathbf{v}_h \rrbracket) \right| &\leq \|\mathbb{C}\| \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket \mathbf{u}_h \rrbracket)\|_{0,E} \|\mathbf{r}_e(\llbracket \mathbf{v}_h \rrbracket)\|_{0,E}.
\end{aligned} \tag{3.15}$$

Adding each term over all elements and using the Cauchy-Schwartz inequality yields (i). The constant  $M$  depends on  $\|\mathbb{C}\|$ ,  $N_e$ , and  $\beta$ , but is independent of  $h$ .

Now we show coercivity, (ii). To simplify the notation, define

$$\|\gamma\|_{0,E,\mathbb{C}}^2 = \int_E \gamma \cdot \mathbb{C} \cdot \gamma \quad \forall \gamma \in W_h^s. \tag{3.16}$$

Due to (3.13), we get

$$\begin{aligned}
a_h(\mathbf{u}_h, \mathbf{u}_h) &= \sum_{E \in \mathcal{T}_h} \left( \|\nabla_s \mathbf{u}_h\|_{0,E,\mathbb{C}}^2 + \int_E 2 \nabla_s \mathbf{u}_h \cdot \mathbb{C} \cdot \mathbf{R}(\llbracket \mathbf{u}_h \rrbracket) + \beta \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket \mathbf{u}_h \rrbracket)\|_{0,E,\mathbb{C}}^2 \right) \\
&\geq \sum_{E \in \mathcal{T}_h} \left( (1 - \varepsilon) \|\nabla_s \mathbf{u}_h\|_{0,E,\mathbb{C}}^2 - \frac{1}{\varepsilon} \|\mathbf{R}(\llbracket \mathbf{u}_h \rrbracket)\|_{0,E,\mathbb{C}}^2 + \beta \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket \mathbf{u}_h \rrbracket)\|_{0,E,\mathbb{C}}^2 \right) \\
&\geq \sum_{E \in \mathcal{T}_h} \left( (1 - \varepsilon) \|\nabla_s \mathbf{u}_h\|_{0,E,\mathbb{C}}^2 + \left( \beta - \frac{N_e}{\varepsilon} \right) \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket \mathbf{u}_h \rrbracket)\|_{0,E,\mathbb{C}}^2 \right),
\end{aligned} \tag{3.17}$$

where we used the standard inequality,  $2ab \leq \varepsilon a^2 + b^2/\varepsilon$ , for all  $\varepsilon > 0$ . Any  $\beta > N_e$  guarantees that  $(\beta - N_e/\varepsilon) > 0$  whenever  $N_e/\beta < \varepsilon < 1$ . Since each term is positive, we can invoke (2.2) to deduce (ii) with  $\mu = c(\beta - N_e/\varepsilon) > 0$ .  $\blacksquare$

**Remark 3.6.** As suggested in [7, 8], following the same steps as in the previous proof establishes the continuity and coercivity of the bilinear form given by the left-hand side of equation (2.32), but for any  $\beta > 0$ .



The following lemma is a preliminary to proving convergence of the discrete displacement, first in  $\|\cdot\|_s$ , and subsequently in  $L^2(B)$ .

**Lemma 3.7.** Let  $u \in (H^1(B))^d$  with  $\nabla \cdot (\mathbb{C} \cdot \nabla u) \in (L^2(B))^d$ , and let  $v_h \in V_h$ , then

$$\sum_{E \in \mathcal{T}_h} \int_{\partial E} n_E \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h = \sum_{e \in \mathcal{E}_h} \int_e \{n \cdot \mathbb{C} \cdot \nabla u\} \cdot \llbracket v_h \rrbracket. \quad (3.18)$$

□

*Proof.* The assumed regularity of  $u$  implies that  $n \cdot \mathbb{C} \cdot \nabla u$  is continuous across interelement boundaries (e.g., [27, Theorem 1.3]); that is,  $0 = n \cdot (\mathbb{C} \cdot \nabla u^- - \mathbb{C} \cdot \nabla u^+)$  on any face in  $\mathcal{E}_h^I$ . Therefore

$$\begin{aligned} & \sum_{E \in \mathcal{T}_h} \int_{\partial E} n_E \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\ &= \sum_{e \in \mathcal{E}_h^I} \int_e (-n \cdot (\mathbb{C} \cdot \nabla u^+) \cdot v_h^+ + n \cdot (\mathbb{C} \cdot \nabla u^-) \cdot v_h^-) + \sum_{e \in \mathcal{E}_h^D} \int_e n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\ &= \sum_{e \in \mathcal{E}_h^I} \int_e -\frac{1}{2} (n \cdot (\mathbb{C} \cdot \nabla u^+) + n \cdot (\mathbb{C} \cdot \nabla u^-)) \cdot v_h^+ \\ & \quad + \frac{1}{2} (n \cdot (\mathbb{C} \cdot \nabla u^+) + n \cdot (\mathbb{C} \cdot \nabla u^-)) \cdot v_h^- + \sum_{e \in \mathcal{E}_h^D} \int_e n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\ &= \sum_{e \in \mathcal{E}_h} \int_e \{n \cdot \mathbb{C} \cdot \nabla u\} \cdot \llbracket v_h \rrbracket. \end{aligned} \quad (3.19)$$

■

The next component of the convergence proof is a bound on the approximation error  $\|u - u_I\|_s$  when  $u_I$  is a suitable interpolant of the exact solution  $u$ . Arnold, Brezzi, Cockburn, and Marini [3] note that discontinuous interpolants can be employed if they satisfy a local approximation property summarized in the next theorem.

**Theorem 3.8** (local interpolation-error estimate). For  $v \in (H^{k+1}(E))^d$ , let  $v_I$  be the  $\mathcal{P}_k$ -interpolant of  $v$  on  $E \in (\mathcal{T}_h)$ . There exists  $C > 0$ , independent of  $E \in (\mathcal{T}_h)$  and therefore of  $h$ , such that

$$|v - v_I|_{q,E} \leq Ch^{k+1-q} |v|_{k+1,E}, \quad k+1 \geq q \geq 0, \quad (3.20)$$

provided  $\mathcal{P}_k(E) \subseteq V_h^E \subset H^q(E)$ . □

*Proof.* The proof is given by Ciarlet [10, Theorem 3.1.5]. ■

**Theorem 3.9** (interpolation-error estimate). Let  $u \in (H^m(B))^d$  for some  $m$  such that  $2 \leq m \leq k+1$ , and let  $u_I \in V_h$  be the  $\mathcal{P}_k$ -interpolant of  $u$  over each element in  $\mathcal{T}_h$ . Then the following interpolation inequality holds:

$$\|u - u_I\|_s \leq Ch^{m-1}|u|_{m,B}, \quad (3.21)$$

where  $C > 0$  is a constant depending only on  $d, m$ , and the upper bound on the Lipschitz constant of the boundary for every element  $E \in \mathcal{T}_h$ , but not on  $h$  or the function  $u$ .  $\square$

*Proof.* From the previous theorem, we have

$$\sum_{E \in \mathcal{T}_h} |u - u_I|_{q,E}^2 \leq \sum_{E \in \mathcal{T}_h} Ch^{2m-2q}|u|_{m,E}^2, \quad m \geq q. \quad (3.22)$$

In addition, the trace inequality [16, page 133] together with a scaling argument gives

$$\|u\|_{0,e}^2 \leq C(h^{-1}|u|_{0,E}^2 + h|u|_{1,E}^2) \quad \forall u \in H^1(E), \quad (3.23)$$

where the constant  $C$  depends only on the Lipschitz constant of the boundary of the element, and can be chosen to be the same for all elements in the family of subdivisions  $(\mathcal{T}_h)$  under consideration.

Following [3], the interpolation inequality (3.21) is established using the inequality (3.23), the bound (3.22), and the inverse inequality (3.10). Starting from the definition of  $\|\cdot\|_s$ , the theorem is obtained as follows:

$$\begin{aligned} \|u - u_I\|_s^2 &= \sum_{E \in \mathcal{T}_h} \|\nabla_s(u - u_I)\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u - u_I \rrbracket)\|_{0,B}^2 \\ &\leq \sum_{E \in \mathcal{T}_h} \|\nabla(u - u_I)\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u - u_I \rrbracket)\|_{0,B}^2 \\ &\leq \sum_{E \in \mathcal{T}_h} |u - u_I|_{1,E}^2 + \sum_{e \in \mathcal{E}_h} Ch^{-1} \|\llbracket u - u_I \rrbracket\|_{0,e}^2 \\ &\leq C \sum_{E \in \mathcal{T}_h} h^{2m-2}|u|_{m,E}^2 \\ &\leq Ch^{2m-2}|u|_{m,B}^2. \end{aligned} \quad (3.24)$$

Again, the constant  $C$  is positive and depends only on  $d, m$ , and the upper bound on the Lipschitz constant of the boundary for every element  $E \in \mathcal{T}_h$ , but not on  $h$  or the function  $u$ .  $\blacksquare$

The last ingredient for the convergence proof is an analysis of the consistency error in the bilinear form (2.34) for functions in  $\widehat{V}_h$ .

**Remark 3.10.** The bilinear form (2.35), which coincides with (2.34) on  $V_h$ , is consistent but continuity does not hold with respect to the norm  $\|\cdot\|_s$  for functions in  $\widehat{V}_h$ . However, it can be shown that (2.34) is continuous with respect to a different norm on a different function space in which some additional regularity is requested [3].

**Theorem 3.11** (bound on consistency error). Let  $u$  be the exact solution to (2.1), with  $u \in (H^m(B))^d$  for some  $m$  such that  $2 \leq m \leq k+1$ . Then

$$\left| a_h(u, v_h) - \int_B f \cdot v_h \right| \leq Ch^{m-1} |u|_{m,B} \|v_h\|_s \quad (3.25)$$

for all  $v_h \in \widehat{V}_h$ . In addition, if  $v_h$  is continuous, then

$$a_h(u, v_h) = \int_B f \cdot v_h. \quad (3.26)$$

□

**Proof.** Recall that  $\llbracket u \rrbracket = 0$ , and use the definition of the bilinear form  $a_h(\cdot, \cdot)$ , (2.34), integration by parts, and application of Lemma 3.7 to obtain

$$\begin{aligned} a_h(u, v_h) - \int_B f \cdot v_h &= \sum_{E \in \mathcal{T}_h} \left( \int_E \nabla_s u \cdot \mathbb{C} \cdot \nabla_s v_h + \int_E \nabla_s u \cdot \mathbb{C} \cdot R(\llbracket v_h \rrbracket) \right) - \int_B f \cdot v_h \\ &= \sum_{E \in \mathcal{T}_h} \left( \int_E \nabla u \cdot \mathbb{C} \cdot \nabla v_h + \int_E \nabla_s u \cdot \mathbb{C} \cdot R(\llbracket v_h \rrbracket) - \int_E f \cdot v_h \right) \\ &= \sum_{E \in \mathcal{T}_h} \left( - \int_E (\nabla \cdot (\mathbb{C} \cdot \nabla u) + f) \cdot v_h + \int_{\partial E} n_E \cdot \mathbb{C} \cdot \nabla u \cdot v_h + \int_E \nabla_s u \cdot \mathbb{C} \cdot R(\llbracket v_h \rrbracket) \right) \\ &= \sum_{e \in \mathcal{E}_h} \int_e \llbracket v_h \rrbracket \cdot \{ n \cdot \mathbb{C} \cdot \nabla_s u \} + \sum_{E \in \mathcal{T}_h} \int_E \nabla_s u \cdot \mathbb{C} \cdot R(\llbracket v_h \rrbracket) \\ &= \sum_{e \in \mathcal{E}_h} \int_e \llbracket v_h \rrbracket \cdot \{ n \cdot \mathbb{C} \cdot \nabla_s (u - u_I) \} + \sum_{E \in \mathcal{T}_h} \int_E \nabla_s (u - u_I) \cdot \mathbb{C} \cdot R(\llbracket v_h \rrbracket). \end{aligned} \quad (3.27)$$

The last line comes from the definition, (2.23), of  $R(v)$  with  $\gamma_h = -\mathbb{C} \cdot \nabla u_I$  and  $v = \llbracket v_h \rrbracket$ . We have also used the symmetry of  $\mathbb{C}$  to interchange gradients and symmetric gradients. Note that if  $v_h$  is continuous, then  $\llbracket v_h \rrbracket = 0$ , and the bilinear form is consistent.

To complete the proof, we bound the consistency error as follows. Using the trace inequality, (3.23), for  $u - u_I \in H^2(E)$ , and (3.9), we have

$$\begin{aligned}
& \left| \sum_{e \in \mathcal{E}_h} \int_e \llbracket v_h \rrbracket \cdot \{n \cdot \mathbb{C} \cdot \nabla(u - u_I)\} \right| \\
& \leq C \sum_{E \in \mathcal{T}_h} \sum_{e \subset \partial E} h^{1/2} \|r_e(\llbracket v_h \rrbracket)\|_{0,B} \|\nabla(u - u_I)\|_{0,B} \\
& \leq C \sum_{E \in \mathcal{T}_h} \sum_{e \subset \partial E} h^{1/2} \|r_e(\llbracket v_h \rrbracket)\|_{0,B} \left( h^{-1} |u - u_I|_{1,E}^2 + h |u - u_I|_{2,E}^2 \right)^{1/2} \quad (3.28) \\
& \leq C \sum_{E \in \mathcal{T}_h} \sum_{e \subset \partial E} \|r_e(\llbracket v_h \rrbracket)\|_{0,B} \left( |u - u_I|_{1,E} + h |u - u_I|_{2,E} \right) \\
& \leq Ch^{m-1} |u|_{m,B} \|v_h\|_s.
\end{aligned}$$

Moreover,

$$\begin{aligned}
& \left| \sum_{E \in \mathcal{T}_h} \int_E \nabla_s(u - u_I) \cdot \mathbb{C} \cdot R(v_h) \right| \leq C \sum_{E \in \mathcal{T}_h} \|\nabla_s(u - u_I)\|_{0,E} \|R(v_h)\|_{0,E} \\
& \leq C \sum_{E \in \mathcal{T}_h} |u - u_I|_{1,E} \sum_{e \subset \partial E} \|r_e(v_h)\|_{0,E} \quad (3.29) \\
& \leq Ch^{m-1} |u|_{m,B} \|v_h\|_s. \quad \blacksquare
\end{aligned}$$

**Corollary 3.12.** Let  $u$  be the exact solution to (2.1), with  $u \in (H^m(B))^d$  for some  $m$  such that  $2 \leq m \leq k+1$ , and let  $u_h$  be the solution of (2.33). Then

$$|a_h(u - u_h, v_h)| \leq Ch^{m-1} |u|_{m,B} \|v_h\|_s \quad (3.30)$$

for all  $v_h \in V_h$ . In addition, if  $v_h$  is continuous, then

$$a_h(u - u_h, v_h) = 0. \quad (3.31) \quad \square$$

**Proof.** Note that

$$\begin{aligned}
a(u - u_h, v_h) &= a_h(u, v_h) - a_h(u_h, v_h) \\
&= a_h(u, v_h) - \int_B f \cdot v_h \quad (3.32)
\end{aligned}$$

for all  $v_h \in V_h$ . The conclusion follows from the previous theorem.  $\blacksquare$

Remark 3.13. Equation (3.31) expresses a Galerkin orthogonality which follows from consistency.

At this point we have gathered all the necessary ingredients to prove convergence of the discrete solutions in  $||| \cdot |||_s$  and  $\| \cdot \|_{0,B}$ , which is the content of the next two theorems.

**Theorem 3.14** (convergence in the mesh-dependent norm  $||| \cdot |||_s$ ). Let  $u$  be the exact solution to (2.1), with  $u \in (H^m(B))^d$  for some  $m$  such that  $2 \leq m \leq k+1$ , and let  $u_h$  be the solution of (2.33), then the following estimate holds:

$$|||u - u_h|||_s \leq Ch^{m-1}|u|_{m,B}, \quad (3.33)$$

where  $C$  is a positive constant independent of  $h$ . □

Proof. From Proposition 3.5, we have

$$\begin{aligned} \mu |||u_I - u_h|||_s^2 &\leq a_h(u_I - u_h, u_I - u_h) \\ &= a_h(u_I - u, u_I - u_h) + a_h(u - u_h, u_I - u_h) \\ &\leq M |||u_I - u_h|||_s |||u_I - u|||_s + |a_h(u - u_h, u_I - u_h)|. \end{aligned} \quad (3.34)$$

The last term in the above equation is due to the general lack of Galerkin orthogonality, but it is appropriately bounded using Corollary 3.12. Therefore,

$$\mu |||u_I - u_h|||_s^2 \leq M |||u_I - u_h|||_s |||u_I - u|||_s + Ch^{m-1}|u|_{m,B} |||u_I - u_h|||_s. \quad (3.35)$$

The conclusion (3.33) follows using (3.21) above. ■

**Theorem 3.15** (convergence in  $L^2(B)$ ). Let  $u$  be the exact solution of (2.1), with  $u \in (H^m(B))^d$  for some  $m$  such that  $2 \leq m \leq k+1$ , and let  $u_h$  be the solution of (2.33), then the following estimate holds:

$$\|u - u_h\|_{0,B} \leq Ch^m|u|_{m,B}. \quad (3.36)$$

□

Proof. The proof follows a standard duality argument. Consider the adjoint problem. Find  $w \in (H^2(B))^d$  such that

$$\begin{aligned} -\nabla \cdot (\mathbb{C} \cdot \nabla_s w) &= u - u_h \quad \text{in } B, \\ w &= 0 \quad \text{on } \partial B. \end{aligned} \quad (3.37)$$

Since  $u - u_h \in (L_2(B))^d$ , the following standard elliptic regularity estimate holds (see, e.g., [19]):

$$\|w\|_{2,B} \leq C \|u - u_h\|_{0,B} \quad (3.38)$$

for some constant  $C > 0$ . For  $w \in H^2(B)$ , let  $w_I \in V_h$  be the continuous, piecewise linear interpolant of  $w$  over each element. Apply (3.25) to  $w$ , with  $v_h = u - u_h$  and  $m = 2$ ,

$$\left| a_h(w, u - u_h) - \int_B (u - u_h)^2 \right| \leq Ch|w|_{2,B} \|u - u_h\|_s \quad (3.39)$$

or

$$\|u - u_h\|_{0,B}^2 \leq |a_h(w, u - u_h)| + Ch|w|_{2,B} \|u - u_h\|_s. \quad (3.40)$$

Since  $w_I$  is continuous, Corollary 3.12 shows  $a_h(w_I, u - u_h) = 0$ . This fact and continuity of the bilinear form, Proposition 3.5, allow us to conclude that

$$\begin{aligned} \|u - u_h\|_{0,B}^2 &\leq |a_h(w - w_I, u - u_h)| + Ch|w|_{2,B} \|u - u_h\|_s \\ &\leq M \|w - w_I\|_s \|u - u_h\|_s + Ch|w|_{2,B} \|u - u_h\|_s \\ &\leq Ch|w|_{2,B} \|u - u_h\|_s, \end{aligned} \quad (3.41)$$

where we have used Theorem 3.9 for the interpolation error estimate  $\|w - w_I\|_s$ . The conclusion of the proof follows from (3.38) and Theorem 3.14.  $\blacksquare$

**Corollary 3.16** (convergence of the stress in  $L^2(B)$ ). Let  $\sigma$  be the exact solution with components in  $H^{m-1}(B)$  for some  $m$  such that  $2 \leq m \leq k + 1$ , and let  $\sigma_h$  be given by (2.25), then the following estimate holds:

$$\|\sigma - \sigma_h\|_{0,B} \leq Ch^{m-1} |u|_{m,B}. \quad (3.42)$$

$\square$

**Proof.** For the exact solution, the displacement is continuous,  $\llbracket u \rrbracket = 0$ , which implies  $R(\llbracket u \rrbracket) = 0$ . So we can write  $\sigma = \mathbb{C} \cdot (\nabla_s u + R(\llbracket u \rrbracket))$ . Therefore,

$$\sigma - \sigma_h = \mathbb{C} \cdot \nabla_s (u - u_h) + \mathbb{C} \cdot R(\llbracket u - u_h \rrbracket). \quad (3.43)$$

It follows that

$$\begin{aligned}
\|\sigma - \sigma_h\|_{0,B}^2 &= \sum_{E \in \mathcal{T}_h} \|\sigma - \sigma_h\|_{0,E}^2 \\
&= \sum_{E \in \mathcal{T}_h} \|\mathbb{C} \cdot \nabla_s(u - u_h) + \mathbb{C} \cdot \mathbf{R}(\llbracket u - u_h \rrbracket)\|_{0,E}^2 \\
&\leq \sum_{E \in \mathcal{T}_h} C \left( \|\nabla_s(u - u_h)\|_{0,E}^2 + N_e \sum_{e \subset \partial E} \|\mathbf{r}_e(\llbracket u - u_h \rrbracket)\|_{0,E}^2 \right) \\
&\leq C \|u - u_h\|_s^2 \leq Ch^{2m-2} |u|_{m,B}^2.
\end{aligned} \tag{3.44}$$

■

Note that this corollary gives  $L^2(B)$  convergence of the stress, even though no such result holds for the strain. This discrepancy is possible because the discrete stress is given by (2.25), and is not, in general, proportional to the strain.

Remark 3.17. Again, as suggested in [7, 8], it can also be proved that the same error estimates hold for the problem directly derived from the variational principle, equation (2.32).

### 3.2 The natural (suboptimal but mesh-independent) BD-estimate

Possible discontinuities in the displacement across element boundaries naturally lead to seeking error estimates in  $BD(B)$ , the space of bounded deformations. This space is defined as the set of functions  $u \in L^1(B)$  whose symmetric part of the distributional derivative  $Du$ ,  $\mathcal{E}(Du) = (1/2)(Du + Du^T)$ , is a matrix-valued bounded Radon measure.

For a function  $u \in BD(B)$ , let  $\|\mathcal{E}(Du)\|(B)$  denote the total symmetric variation measure of  $Du$ . A general Poincaré-type estimate for BD-functions holds in the following form.

**Theorem 3.18** (Poincaré inequality for BD). Let  $B \subset \mathbb{R}^d$  be a bounded domain with Lipschitz boundary. Then there exists  $C > 0$  such that for all  $u \in BD(B)$ ,  $u|_{\partial B} = 0$ ,

$$\|u\|_{L^1(B)} \leq C \|\mathcal{E}(Du)\|(B), \tag{3.45}$$

where  $u|_{\partial B}$  denotes the generalized trace. □

Proof. The proof is given by Temam [29, Remark II.2.5, page 189]. ■

**Theorem 3.19** (natural BD estimate). There exists  $C > 0$  such that for all  $\mathbf{u} \in \widehat{\mathbf{V}}$ ,

$$\|\mathbf{u}\|_{\text{BD}(\mathbf{B})} \leq C \|\mathbf{u}\|_s, \quad (3.46)$$

with  $C$  independent of  $h$ . □

*Proof.* Recall the definition of the BD norm

$$\|\mathbf{u}\|_{\text{BD}(\mathbf{B})} = \|\mathbf{u}\|_{L^1(\mathbf{B})} + \|\mathcal{E}(\mathbf{D}\mathbf{u})\|(\mathbf{B}), \quad (3.47)$$

where

$$\|\mathcal{E}(\mathbf{D}\mathbf{u})\|(\mathbf{B}) = \sup \left\{ \int_{\mathbf{B}} \mathbf{u} \cdot (\nabla \cdot (\Psi^T + \Psi)) : \Psi \in C_0^1(\mathbf{B}, \mathbb{R}^{d \times d}), \|\Psi\|_{L^\infty} \leq 1 \right\}. \quad (3.48)$$

The proof continues mutatis mutandis as in [Theorem 3.26](#). ■

Using the estimate (3.46) for the difference  $\mathbf{u} - \mathbf{u}_h$  together with [Theorem 3.9](#) shows that convergence of the method is immediately strengthened from the  $\|\cdot\|_s$ -norm to a mesh-independent estimate in the space  $\text{BD}(\mathbf{B})$ . It is clear that any “optimal” estimate in the symmetric norm, derived under less smoothness assumptions on the underlying continuous problem [26], translates into a corresponding “optimal” mesh-independent BD estimate. It is worth remarking that the derivation of the BD estimate does not make use of [Theorem 3.15](#) that additionally establishes convergence of the discrete solutions in  $L^2(\mathbf{B})$ .

The occurrence of the space BD is, strictly speaking, an artifact of the linearized treatment, where only the symmetrized infinitesimal strains  $\varepsilon(\nabla \mathbf{u})$  appear. Since this BD estimate does not control the antisymmetric part of the displacement gradient, we are interested in obtaining convergence in the space  $\text{BV}(\mathbf{B})$ . However, since  $\text{BV}(\mathbf{B})$  is strictly smaller than  $\text{BD}(\mathbf{B})$ , there is no obvious way to proceed directly from the BD estimate to a BV estimate. Instead, we will first strengthen [Theorem 3.14](#) to the  $\|\cdot\|$ -norm. Note that for a given mesh size  $h > 0$ , given the finite dimensionality of  $\mathbf{V}_h$  and the fact that both  $\|\cdot\|$  and  $\|\cdot\|_s$  are norms in  $\mathbf{V}_h$ , we have, for  $\mathbf{u}_h \in \mathbf{V}_h$ ,

$$\|\mathbf{u}_h\|_{\text{BD}} \leq \|\mathbf{u}_h\|_{\text{BV}} \leq C \|\mathbf{u}_h\| \leq c(h) \|\mathbf{u}_h\|_s, \quad (3.49)$$



where the estimate  $\|u_h\|_{BV} \leq C\|u_h\|$  is obtained in [Theorem 3.26](#). However,  $c(h)$  may not be bounded from below away from zero for all  $h > 0$ . The failure to obtain a mesh-independent estimate between  $\|u_h\|_s$  and  $\|u_h\|$  is a manifestation of the possible lack of a discrete Korn's first inequality for nonconforming meshes [\[17\]](#). In order to obtain convergence in the  $\|\cdot\|$ -norm, followed by a BV-estimate, and then convergence in BV, we first establish a generalized version of Korn's second inequality at the element level.

### 3.3 Korn's second inequality for the subdivision

In this section, we investigate an analog to Korn's second inequality at the element level, independent of the element shape and size. The derivation of this inequality relies heavily on how Korn's second inequality scales under uniform contractions. We set  $SL(d, \mathbb{R}) = \{X \in \mathbb{R}^{d \times d} \mid \det X = 1\}$ .

**Theorem 3.20** (Korn's second inequality under distortion). Assume that  $\Omega \subset \mathbb{R}^d$  is a bounded (reference) domain with Lipschitz boundary  $\partial\Omega$  and let  $\mathcal{M} = \{X \in SL(d, \mathbb{R}) : \|X\| \leq K\}$ , for some  $K > 0$ . For  $F \in \mathcal{M}$  define  $\Omega_\xi = F(\Omega)$ . Then there exists  $C > 0$  such that for all  $F \in \mathcal{M}$ ,  $u \in H^1(\Omega_\xi)$ ,

$$\|\nabla_\xi u^T + \nabla_\xi u\|_{0, \Omega_\xi}^2 + \|u\|_{0, \Omega_\xi}^2 \geq C\|u\|_{1, \Omega_\xi}^2. \quad (3.50)$$

□

**Proof.** We first translate the statement to the fixed reference domain  $\Omega$ . The affine transformation  $\xi = F(x)$  together with the definition  $u(\xi) = u(F(x)) = \tilde{u}(x)$  and  $\det F = 1$  leads to

$$\int_{\Omega_\xi} \|\nabla_\xi u^T + \nabla_\xi u\|^2 + \|u\|^2 = \int_{\Omega} \|F^{-T} \nabla \tilde{u}^T + \nabla \tilde{u} F^{-1}\|^2 + \|\tilde{u}\|^2. \quad (3.51)$$

We proceed by contradiction. Assume, without loss of generality, that there exists a sequence  $\{\tilde{u}_n\} \in H^1(\Omega)$  with  $\|\tilde{u}_n\|_{1, \Omega} = 1$  and a sequence  $F_n \in \mathcal{M}$  such that

$$\|F_n^{-T} \nabla \tilde{u}_n^T + \nabla \tilde{u}_n F_n^{-1}\|_{0, \Omega}^2 + \|\tilde{u}_n\|_{0, \Omega}^2 \leq \frac{1}{n} \|\tilde{u}_n\|_{1, \Omega}^2 = \frac{1}{n}. \quad (3.52)$$

Since  $F_n$  is bounded, we may extract a subsequence which converges strongly to  $\hat{F} \in \mathcal{M}$  by Bolzano-Weierstrass. It is readily seen by continuity and the boundedness of  $\tilde{u}_n$  that

$$\|\hat{F}^{-T} \nabla \tilde{u}_n^T + \nabla \tilde{u}_n \hat{F}^{-1}\|_{0, \Omega}^2 + \|\tilde{u}_n\|_{0, \Omega}^2 \rightarrow 0. \quad (3.53)$$

Thus  $\tilde{u}_n$  is a minimizing sequence. For fixed  $\hat{F}$ , the quadratic expression is uniformly positive (generalized Korn's second inequality, see [20]) such that

$$\|\hat{F}^{-T} \nabla \tilde{u}_n^T + \nabla \tilde{u}_n \hat{F}^{-1}\|_{0,\Omega}^2 + \|\tilde{u}_n\|_{0,\Omega}^2 \geq C(\hat{F}) \|\tilde{u}_n\|_{1,\Omega}^2 \quad (3.54)$$

for some  $C > 0$ , contradicting  $\|\tilde{u}_n\|_{1,\Omega} = 1$ .  $\blacksquare$

**Theorem 3.21** (Korn's second inequality under scaling). Let  $\Omega \subset \mathbb{R}^d$  be a bounded domain with Lipschitz boundary  $\partial\Omega$  and, without loss of generality,  $|\Omega| = 1$ . Consider the scaled domain  $\Omega_h = \{hx : x \in \Omega\}$ ,  $h > 0$ . Then there exists  $C(\Omega) > 0$  such that for all  $u \in H^1(\Omega_h)$ ,

$$\|\nabla u^T + \nabla u\|_{0,\Omega_h}^2 + \frac{1}{|\Omega_h|^{2/d}} \|u\|_{0,\Omega_h}^2 \geq C(\Omega) \left( \|\nabla u\|_{0,\Omega_h}^2 + \frac{1}{|\Omega_h|^{2/d}} \|u\|_{0,\Omega_h}^2 \right), \quad (3.55)$$

where the constant  $C(\Omega)$  is independent of  $h > 0$  and coincides with the constant in Korn's second inequality for  $\Omega$ .  $\square$

*Proof.* Let  $\tilde{u} \in H^1(\Omega)$ . From Korn's second inequality (see, e.g., [20]), we get

$$\|\nabla \tilde{u}^T + \nabla \tilde{u}\|_{0,\Omega}^2 + \|\tilde{u}\|_{0,\Omega}^2 \geq C(\Omega) (\|\nabla \tilde{u}\|_{0,\Omega}^2 + \|\tilde{u}\|_{0,\Omega}^2). \quad (3.56)$$

Expressing every term with respect to the down-scaled  $\Omega_h$ , where  $\tilde{u}(x) = u(hx)$ , and noticing that  $|\Omega_h| = h^d$ , we get

$$\begin{aligned} \frac{1}{h^{d-2}} \|\nabla u^T + \nabla u\|_{0,\Omega_h}^2 + \frac{1}{h^d} \|u\|_{0,\Omega_h}^2 \\ \geq C(\Omega) \left( \frac{1}{h^{d-2}} \|\nabla u\|_{0,\Omega_h}^2 + \frac{1}{h^d} \|u\|_{0,\Omega_h}^2 \right), \end{aligned} \quad (3.57)$$

from which we deduce the required result. Note that  $C(\Omega)$  is just the constant in Korn's second inequality.  $\blacksquare$

**Corollary 3.22** (uniformity in  $(\mathcal{T}_h)$ ). Let  $\hat{E}$  be the reference element for an element  $E \in (\mathcal{T}_h)$  as defined in Section 2. Without loss of generality, take  $|\hat{E}| = 1$ . Then there exists  $C > 0$  such that for all  $E \in (\mathcal{T}_h)$ ,  $u \in H^1(E)$ ,

$$\|\nabla u^T + \nabla u\|_{0,E}^2 + \frac{1}{|E|^{2/d}} \|u\|_{0,E}^2 \geq C \left( \|\nabla u\|_{0,E}^2 + \frac{1}{|E|^{2/d}} \|u\|_{0,E}^2 \right). \quad (3.58)$$

$\square$

Proof. Let  $F$  be an affine transformation such that  $E = F(\widehat{E})$ . Decompose  $F = F_V \cdot \tilde{F}$  into its isochoric and volumetric part, where  $F_V = (\det F)^{1/d} I$ ,  $I$  is the second-order identity tensor, and  $\tilde{F} = F/(\det F)^{1/d}$ . Note that  $|E| = \det F$ . Using [11, Theorem 3.1.3, page 120] and the quasi-uniformity of the subdivision, we have that

$$\|\tilde{F}\| = \frac{\|F\|}{(\det F)^{1/d}} \leq \frac{h}{\hat{\rho}} \frac{1}{|E|^{1/d}} \leq \frac{C}{\hat{\rho}}, \quad (3.59)$$

where  $\hat{\rho}$  is the diameter of the largest ball contained in  $\widehat{E}$  and  $C$  is independent of  $E$ . Therefore, by Theorem 3.20, we can state Korn's second inequality for each domain  $\tilde{F}(\widehat{E})$  in the subdivision with the same constant  $C > 0$ . The corollary then follows from Theorem 3.21. ■

### 3.4 Convergence in $||| \cdot |||$

We can now obtain convergence of the sequence of discrete solutions in the mesh-dependent norm  $||| \cdot |||$  using our generalized Korn's second inequality for the subdivision.

**Theorem 3.23** (convergence in the mesh-dependent norm  $||| \cdot |||$ ). Let  $(v_h) \subset \widehat{V}$  be a sequence such that  $|||v_h|||_s \leq Ch^{m-1}$  and  $\|v_h\|_{0,B} \leq Ch^m$  for  $h \downarrow 0$ . Then

$$|||v_h||| \leq Ch^{m-1} \quad (3.60)$$

for some  $C > 0$  independent of  $h$ . □

Proof. Use Corollary 3.22 and sum over the elements to obtain the estimate

$$\sum_{E \in \mathcal{T}_h} \left( \|\nabla v_h^T + \nabla v_h\|_{0,E}^2 + \frac{1}{|E|^{2/d}} \|v_h\|_{0,E}^2 \right) \geq C \sum_{E \in \mathcal{T}_h} \left( \|\nabla v_h\|_{0,E}^2 + \frac{1}{|E|^{2/d}} \|v_h\|_{0,E}^2 \right) \quad (3.61)$$

which, in light of equation (2.7), can be weakened to

$$\sum_{E \in \mathcal{T}_h} \left( \|\nabla v_h^T + \nabla v_h\|_{0,E}^2 + \frac{1}{h^2} \|v_h\|_{0,E}^2 \right) \geq C \sum_{E \in \mathcal{T}_h} \left( \|\nabla v_h\|_{0,E}^2 + \frac{1}{h^2} \|v_h\|_{0,E}^2 \right), \quad (3.62)$$

where  $C$  is independent of  $h > 0$ . Without loss of generality, assume  $0 < C \leq 1$ . Adding the specific jump contribution over the faces of each element shows that

$$\begin{aligned} & \sum_{E \in \mathcal{T}_h} \|\nabla v_h^T + \nabla v_h\|_{0,E}^2 + \frac{1}{h^2} \|v_h\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket v_h \rrbracket)\|_{0,B}^2 \\ & \geq C \sum_{E \in \mathcal{T}_h} \left( \|\nabla v_h\|_{0,E}^2 + \frac{1}{h^2} \|v_h\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket v_h \rrbracket)\|_{0,B}^2 \right) \end{aligned} \quad (3.63)$$

or

$$\|v_h\|_s^2 + \frac{1}{h^2} \sum_{E \in \mathcal{T}_h} \|v_h\|_{0,E}^2 \geq C \left( \|v_h\|^2 + \frac{1}{h^2} \sum_{E \in \mathcal{T}_h} \|v_h\|_{0,E}^2 \right), \quad (3.64)$$

where, again,  $C > 0$  is independent of  $h > 0$ . Thus

$$\|v_h\|_s^2 + \frac{1}{h^2} \|v_h\|_{0,B}^2 \geq C \left( \|v_h\|^2 + \frac{1}{h^2} \|v_h\|_{0,B}^2 \right) \geq C \|v_h\|^2. \quad (3.65)$$

Using the convergence of  $(v_h)$  and equation (3.65), we obtain

$$\|v_h\|^2 \leq C \left( h^{2m-2} + \frac{1}{h^2} h^{2m} \right) = Ch^{2m-2}, \quad (3.66)$$

which completes the theorem. ■

**Remark 3.24.** As is evident from the statement of [Theorem 3.23](#), the convergence in  $\|\cdot\|$  can only be shown for sequences converging in both  $\|\cdot\|_s$  and  $\|\cdot\|_{L^2(B)}$  with specific rates in  $h$ , which we have established for  $v_h = u - u_h$  under appropriate hypotheses. In general, for solutions of the continuous problem with less regularity, one might not have such knowledge.

### 3.5 Convergence in BV

We prove that the mesh-dependent norm  $\|\cdot\|$  estimates the BV norm on  $\widehat{V} = V_h + (H_0^1(B))^d$  and as a result, we obtain convergence in BV. Recall that  $BV(B)$  is the space of functions  $u \in L^1(B)$  such that the distributional derivative  $Du$  is a matrix-valued bounded Radon measure.

For a function  $u \in BV(B)$ ,  $\|Du\|(B)$  denotes the total variation measure of  $Du$ . A general Poincaré-type estimate for BV-functions holds in the following form.

**Theorem 3.25** (Poincaré inequality for BV). There exists  $C > 0$  such that for all  $u \in BV(\mathbb{R}^d)$ ,

$$\|u\|_{L^{d/(d-1)}(\mathbb{R}^d)} \leq C \|Du\|(\mathbb{R}^d). \quad (3.67)$$

□

Proof. Evans and Gariepy [16, Theorem 1, page 189]. ■

**Theorem 3.26** (natural BV estimate). There exists  $C > 0$  such that for all  $u \in \widehat{V}$ ,

$$\|u\|_{BV} \leq C \|u\|, \quad (3.68)$$

with  $C$  independent of  $h$ . □

Proof. Recall the definition of the BV norm

$$\|u\|_{BV(B)} = \|u\|_{L^1(B)} + \|Du\|(B), \quad (3.69)$$

where

$$\|Du\|(B) = \sup \left\{ \int_B u \cdot (\nabla \cdot \Psi) : \Psi \in C_0^1(B, \mathbb{R}^{d \times d}), \|\Psi\|_{L^\infty} \leq 1 \right\}. \quad (3.70)$$

First observe that

$$\begin{aligned} \int_B u \cdot (\nabla \cdot \Psi) &= \sum_{E \in \mathcal{T}_h} \int_E u \cdot (\nabla \cdot \Psi) \\ &= \sum_{E \in \mathcal{T}_h} \int_E \nabla \cdot (\Psi \cdot u) - \sum_{E \in \mathcal{T}_h} \int_E \Psi \cdot \nabla u \\ &= \sum_{E \in \mathcal{T}_h} \int_{\partial E} n_E \cdot \Psi \cdot u - \sum_{E \in \mathcal{T}_h} \int_E \Psi \cdot \nabla u \\ &= \sum_{e \in \mathcal{E}_h} \int_e n \cdot \Psi \cdot \llbracket u \rrbracket - \sum_{E \in \mathcal{T}_h} \int_E \Psi \cdot \nabla u. \end{aligned} \quad (3.71)$$

Each term in the two sums may be estimated individually by

$$\begin{aligned} \sup_{\|\Psi\|_{L^\infty} \leq 1} \left[ \int_e n \cdot \Psi \cdot \llbracket u \rrbracket \right] &\leq \int_e \llbracket u \rrbracket \cdot \frac{\llbracket u \rrbracket}{\|\llbracket u \rrbracket\|} \leq \|\llbracket u \rrbracket\|_{L^1(e)}, \\ \sup_{\|\Psi\|_{L^\infty} \leq 1} \left[ - \int_E \Psi \cdot \nabla u \right] &\leq \int_E \frac{\nabla u}{\|\nabla u\|} \cdot \nabla u \leq \|\nabla u\|_{L^1(E)}, \end{aligned} \quad (3.72)$$

which yields the preliminary estimate

$$\|Du\|(B) \leq \sum_{e \in \mathcal{E}_h} \|\llbracket u \rrbracket\|_{L^1(e)} + \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{L^1(E)}. \quad (3.73)$$

Applying Hölder's inequality to each term in the sum gives

$$\|Du\|(B) \leq \sum_{e \in \mathcal{E}_h} |e|^{1/2} \|\llbracket u \rrbracket\|_{0,e} + \sum_{E \in \mathcal{T}_h} |E|^{1/2} \|\nabla u\|_{0,E}. \quad (3.74)$$

Taking the square of both sides and using Young's inequality leads to

$$\|Du\|^2(B) \leq 2 \left[ \sum_{e \in \mathcal{E}_h} |e|^{1/2} \|\llbracket u \rrbracket\|_{0,e} \right]^2 + 2 \left[ \sum_{E \in \mathcal{T}_h} |E|^{1/2} \|\nabla u\|_{0,E} \right]^2. \quad (3.75)$$

Now we use the Cauchy-Schwartz inequality for the sums in the brackets to show

$$\begin{aligned} \|Du\|^2(B) &\leq 2 \left[ \left( \sum_{e \in \mathcal{E}_h} (|e|^{1/2})^2 \right)^{1/2} \left( \sum_{e \in \mathcal{E}_h} \|\llbracket u \rrbracket\|_{0,e}^2 \right)^{1/2} \right]^2 \\ &\quad + 2 \left[ \left( \sum_{E \in \mathcal{T}_h} (|E|^{1/2})^2 \right)^{1/2} \left( \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2 \right)^{1/2} \right]^2 \\ &\leq 2 \left( \sum_{e \in \mathcal{E}_h} |e| \right) \left( \sum_{e \in \mathcal{E}_h} \|\llbracket u \rrbracket\|_{0,e}^2 \right) + 2 \left( \sum_{E \in \mathcal{T}_h} |E| \right) \left( \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2 \right) \end{aligned} \quad (3.76)$$

which, by [Lemma 3.3](#), implies

$$\begin{aligned} \|Du\|^2(B) &\leq 2 \left( \sum_{e \in \mathcal{E}_h} |e| \right) \left( Ch \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u \rrbracket)\|_{0,B}^2 \right) + 2|B| \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2 \\ &\leq 2C \left[ \left( \sum_{e \in \mathcal{E}_h} |e|h \right) \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u \rrbracket)\|_{0,B}^2 \right] + 2|B| \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2, \end{aligned} \quad (3.77)$$

with  $C$  independent of  $h$ . From (2.8),

$$\begin{aligned} \|Du\|^2(B) &\leq 2C \left[ \sum_{e \in \mathcal{E}_h} |E| \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u \rrbracket)\|_{0,B}^2 \right] + 2|B| \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2 \\ &\leq C|B| \left[ \sum_{e \in \mathcal{E}_h} \|r_e(\llbracket u \rrbracket)\|_{0,B}^2 + \sum_{E \in \mathcal{T}_h} \|\nabla u\|_{0,E}^2 \right] \\ &\leq C|B| \|u\|^2. \end{aligned} \quad (3.78)$$

By hypothesis,  $u \in V_h + (H_0^1(B))^d$ ; this implies  $u \in BV(B)$  since  $u \in L^2(B)$  and  $\|Du\|(B)$  is bounded by  $\|u\|$ . We may extend  $u$  to a function  $\tilde{u}$  on all of  $\mathbb{R}^d$  by setting  $u$  to zero outside of  $B$ . From [16, Theorem 1, page 183], we have the equivalence

$$\|D\tilde{u}\|(\mathbb{R}^d) = \|Du\|(B). \quad (3.79)$$

Thus, by applying the Poincaré inequality for  $BV$ , Theorem 3.25, we obtain

$$\|u\|_{L^{d/(d-1)}(B)} = \|\tilde{u}\|_{L^{d/(d-1)}(\mathbb{R}^d)} \leq C \|D\tilde{u}\|(\mathbb{R}^d) = C \|Du\|(B) \leq C \|u\| \quad (3.80)$$

with  $C > 0$  independent of  $h$ . This estimate is necessary since the mesh-dependent norm  $\|\cdot\|$  does not contain a contribution of the form  $\|u\|_{L^2(B)}$ . ■

**Corollary 3.27** (optimal mesh-independent estimate). Let  $(v_h) \subset \hat{V}$  be a sequence such that  $\|v_h\|_s \leq Ch^{m-1}$  and  $\|v_h\|_{0,B} \leq Ch^m$  for  $h \downarrow 0$ . Then

$$\|v_h\|_{BV} \leq Ch^{m-1}. \quad (3.81)$$

□

Proof. Apply Theorem 3.23 together with Theorem 3.26. ■

#### 4 Final remarks

Optimal convergence of a stabilized DG method for linear elasticity with Dirichlet boundary conditions has been established in the mesh-independent  $BV$  norm. Unlike interior penalty methods, the stabilization term contains a constant factor  $\beta > N_e$  that is easy to determine for a given discretization. The finite-element spaces composed of piecewise polynomial functions over the elements are also easy to implement. In future work, we will explore the numerical properties of the method and its extensions to finite elasticity, elastoplasticity, and fracture.

## Acknowledgments

Patrizio Neff and Deborah Sulsky acknowledge the kind hospitality of the Graduate Aeronautical Laboratories during their visits. We thank Donatella Marini, Ilaria Perugia, and Dominik Schötzau for comments on an earlier draft of this paper.

## References

- [1] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), no. 4, 742–760.
- [2] D. N. Arnold, F. Brezzi, B. Cockburn, and D. Marini, *Discontinuous Galerkin methods for elliptic problems*, Discontinuous Galerkin Methods: Theory, Computation and Applications (Newport, RI, 1999) (B. Cockburn, G. E. Karniadakis, and C.-W. Shu, eds.), Lect. Notes Comput. Sci. Eng., vol. 11, Springer, Berlin, 2000, pp. 89–101.
- [3] ———, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2002), no. 5, 1749–1779.
- [4] G. A. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp. **31** (1977), no. 137, 45–59.
- [5] F. Bassi and S. Rebay, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys. **131** (1997), no. 2, 267–279.
- [6] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 1994.
- [7] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo, *Discontinuous Galerkin approximations for elliptic problems*, Numer. Methods Partial Differential Equations **16** (2000), no. 4, 365–378.
- [8] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo, *Discontinuous finite elements for diffusion problems*, Atti Convegno in Onore di F. Brioschi (Milano, 1997), Istituto Lombardo, Accademia di Scienze e Lettere, Milano, 2001, pp. 197–217.
- [9] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal. **38** (2000), no. 5, 1676–1706.
- [10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing, Amsterdam, 1978.
- [11] ———, *Mathematical Elasticity: Three-Dimensional Elasticity*, Studies in Mathematics and Its Applications, vol. 20, North-Holland Publishing, Amsterdam, 1988.
- [12] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal. **35** (1998), no. 6, 2440–2463.
- [13] ———, *Runge-Kutta discontinuous Galerkin methods for convection-dominated problems*, J. Sci. Comput. **16** (2001), no. 3, 173–261.



- [14] J. Douglas Jr.,  $H^1$ -Galerkin methods for a nonlinear Dirichlet problem, Mathematical Aspects of Finite Element Methods (Proc. Conf., Consiglio Naz. delle Ricerche (CNR), Rome, 1975), Lecture Notes in Math., vol. 606, Springer, Berlin, 1977, pp. 64–86.
- [15] J. Douglas Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, Computing Methods in Applied Sciences (Second Internat. Sympos., Versailles, 1975), Lecture Notes in Phys., vol. 58, Springer, Berlin, 1976, pp. 207–216.
- [16] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Florida, 1992.
- [17] R. S. Falk, Nonconforming finite element methods for the equations of linear elasticity, Math. Comp. **57** (1991), no. 196, 529–550.
- [18] P. Houston, C. Schwab, and E. Süli, Discontinuous hp-finite element methods for advection-diffusion-reaction problems, SIAM J. Numer. Anal. **39** (2002), no. 6, 2133–2163.
- [19] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity*, Prentice Hall, New York, 1983, reprinted by Dover Publications, New York, 1994.
- [20] P. Neff, On Korn's first inequality with non-constant coefficients, Proc. Roy. Soc. Edinburgh Sect. A **132** (2002), no. 1, 221–243.
- [21] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind [On a variational principle for solving Dirichlet problems less boundary conditions using subspaces], Abh. Math. Sem. Univ. Hamburg **36** (1971), 9–15 (German).
- [22] J. T. Oden, I. Babuška, and C. E. Baumann, A discontinuous hp finite element method for diffusion problems, J. Comput. Phys. **146** (1998), no. 2, 491–519.
- [23] M. Ortiz and A. Pandolfi, Finite-deformation irreversible cohesive elements for three-dimensional crack propagation analysis, Internat. J. Numer. Methods Engrg. **44** (1999), no. 9, 1267–1282.
- [24] P. Percell and M. F. Wheeler, A local residual finite element procedure for elliptic equations, SIAM J. Numer. Anal. **15** (1978), no. 4, 705–714.
- [25] T. H. H. Pian and P. Tong, Basis of finite element methods for solid continua, Internat. J. Numer. Methods Engrg. **1** (1969), 3–28.
- [26] B. Rivière and M. F. Wheeler, *Optimal Error Estimates for Discontinuous Galerkin Methods Applied to Linear Elasticity Problems*, Tech. Report 00-30, Texas Institute for Computational and Applied Mathematics, 2000.
- [27] J. E. Roberts and J.-M. Thomas, Mixed and hybrid methods, Handbook of Numerical Analysis, Vol. II (P. G. Ciarlet and J.-L. Lions, eds.), North-Holland Publishing, Amsterdam, 1991, pp. 523–639.
- [28] V. Ruas, Circumventing discrete Korn's inequalities in convergence analyses of nonconforming finite element approximations of vector fields, Z. Angew. Math. Mech. **76** (1996), no. 8, 483–484.
- [29] R. Temam, *Problèmes Mathématiques en Plasticité* [Mathematical Problems in Plasticity], Méthodes Mathématiques de l'Informatique, vol. 12, Gauthier-Villars, Montrouge, 1983.
- [30] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM J. Numer. Anal. **15** (1978), no. 1, 152–161.

Adrian Lew: Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA

E-mail address: [lewa@stanford.edu](mailto:lewa@stanford.edu)

Patrizio Neff: Fachbereich Mathematik, Technische Universität Darmstadt, 64289 Darmstadt, Germany

E-mail address: [neff@mathematik.tu-darmstadt.de](mailto:neff@mathematik.tu-darmstadt.de)

Deborah Sulsky: Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131, USA

E-mail address: [sulsky@math.nm.edu](mailto:sulsky@math.nm.edu)

Michael Ortiz: The Graduate Aeronautical Laboratories, California Institute of Technology, Pasadena, CA 90025, USA

E-mail address: [ortiz@aero.caltech.edu](mailto:ortiz@aero.caltech.edu)