OPTIMAL BV ESTIMATES FOR A DISCONTINUOUS GALERKIN METHOD IN LINEAR ELASTICITY*

ADRIAN LEW †, PATRIZIO NEFF ‡, DEBORAH SULSKY §, AND MICHAEL ORTIZ ¶

Abstract. We analyze a discontinuous Galerkin method for linear elasticity. The discrete formulation derives from the Hellinger-Reissner variational principle with the addition of stabilization terms analogous to those previously considered by others for the Navier-Stokes equations and a scalar Poisson equation. For our formulation, we first obtain convergence in a mesh-dependent norm and in the natural mesh-independent BD norm. We then prove a generalization of Korn's second inequality which allows us to strengthen our results to an optimal, mesh-independent BV estimate for the error.

Key words. Discontinuous Galerkin Method, Linear Elasticity, BV, Nonconforming Elements, Korn's inequality

AMS subject classifications. 65N12, 65N15, 65N30

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 coworkers [1]. At the time of their introduction, DG methods were generally called interior penalty methods, and were considered by Baker [2], Douglas [3], and Douglas and Dupont [4] for fourth order problems where C^1 continuity was imposed on C^0 elements. For second order equations, Nitsche [5] appears to have introduced the idea of imposing Dirichlet boundary conditions weakly and of adding stabilization terms to obtain optimal convergence rates. The same idea of penalizing jumps along inter-element faces lead to the interior penalty methods of Percell and Wheeler [6], and Wheeler [7]. Methods for a second order, nonlinear, parabolic equation appeared in Arnold [8].

According to [9], 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 [9]. Recently, however, interest has been rekindled by developments in DG methods for convection-diffusion problems; see, for example, Cockburn and Shu [10, 11], Oden, *et al.* [12], Castillo, *et al.* [13], and Houston, *et al.* [14], where the scalar Poisson equation is analyzed. Bassi and Rebay [15] applied a similar technique for the solution of the Navier-Stokes equations. Brezzi, *et al.* [16, 17] analyzed the method of Bassi and Rebay, for stability and accuracy, as it applies to the scalar Poisson equation. Arnold, *et al.* [9, 18] provided a common framework for all of these methods and showed the interconnections by casting them into the form of the local discontinuous Galerkin method (LDG) 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 [19]. There are

---

*

†GALCIT, California Institute of Technology, Pasadena, CA 90025

‡GALCIT, California Institute of Technology, Pasadena, CA 90025 and FB Mathematik, TU Darmstadt, Germany

§Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131

¶GALCIT, California Institute of Technology, Pasadena, CA 90025

several benefits of such an approach, including the potential for efficient hp-adaptivity, for example, using meshes with hanging nodes with adaptive mesh refinement, and the prospect of rigorously handling problems with discontinuous displacements as arise in the study of fracture. Rivière and Wheeler [20] formulate and analyze a method for linear elasticity based on a generalization of the nonsymmetric interior penalty Galerkin (NIPG) method presented in [12] for the diffusion equation. The resulting bilinear form is non-symmetric. As an alternative, we follow the analysis of Brezzi, *et al.* [16, 17] 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 discontinuous Galerkin 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 article 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 leads to a formulation analogous to the one utilized in Bassi and Rebay [15]. Stabilization terms of the form considered in Brezzi, *et al.* [16, 17] 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 $d = 2, 3$,

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

The body B is assumed to be a bounded, polyhedral domain. The function $u : B \to \mathbb{R}^d$ 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, *i.e.*,

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

for all γ 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, $\nabla_s u = \frac{1}{2}\left(\nabla u + (\nabla u)^\mathsf{T}\right)$. The boundary of the domain, ∂B, is decomposed into two disjoint sets, $\partial_D B$ and $\partial_N B$. The body is acted upon by body forces, $f : B \to \mathbb{R}^d$, and surface tractions, $\bar{T} : \partial_N B \to \mathbb{R}^d$. The displacement, $\bar{u} : \partial_D B \to \mathbb{R}^d$, 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

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

The first equation is the constitutive equation that relates the stress tensor σ 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, σ) 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, $I : (H^{m+1}(B))^d \times (H^m(B))^{d\times d} \to \mathbb{R}$, where

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

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

2.2. The discrete scheme. First, we consider a family of subdivisions $(\mathcal{T}_h)$ of B with $h \downarrow 0$. A subdivision $\mathcal{T}_h$ of B into a finite number of sets E, such that $\bar{B} = \cup_{E\in\mathcal{T}_h} E$ is called admissible in the sense of [21, p. 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, ∂E, of each E is Lipschitz continuous. We assume the family $(\mathcal{T}_h)$ to be quasi-uniform [22, p. 106] so that

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

and

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

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

$$ch^d \leq |E| \leq Ch^d \tag{2.7}$$

for every element $E \in \mathcal{T}_h$ and for 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 [22, p. 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

$$|e|h \leq C|E|, \tag{2.8}$$

for all $h > 0$, and for any face e of any element $E \in (\mathcal{T}_h)$. Even though discontinuous Galerkin 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 ∂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, ∂E, with unit, outward normal denoted by n_E. Define the set of internal faces,

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

the set of Dirichlet faces,

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

and the set of Neumann faces,

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

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

$$\Gamma = \cup_{e \in \mathcal{E}_h} e.$$

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 multi-valued on $\Gamma \setminus \partial B$ and single-valued on ∂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 $T(\Gamma)$ [1].

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.

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

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 elements spaces, such as those constructed from Lagrange simplices of various degrees and some spaces constructed with bilinear quadrilaterals or trilinear bricks.

Remark. 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, $\{\cdot\} : T(\Gamma) \to (L^2(\Gamma))^d$, and the jump operator, $[\![\cdot]\!] : T(\Gamma) \to (L^2(\Gamma))^d$. 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

$$\{v\} = \frac{1}{2}(v^-|_e + v^+|_e) \tag{2.9}$$

and the jump by

$$[\![v]\!] = v^-|_e - v^+|_e. \tag{2.10}$$

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

$$\{v\} = v, \qquad \text{and} \quad [\![\mathrm{v}]\!] = \mathrm{v}; \tag{2.11}$$

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

$$\{v\} = v, \qquad \text{and} \quad [\![\mathrm{v}]\!] = 0. \tag{2.12}$$

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 ∂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

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

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

$$\begin{aligned} I_E = \int_E \Big(\frac{1}{2}\sigma \cdot \mathbb{C}^{-1} \cdot \sigma - \sigma \cdot \nabla_s u + f \cdot u\Big) + \int_{\partial E \setminus \partial B} \frac{1}{2} n_E \cdot \sigma \cdot (u - u^{\text{ext}}) \\ + \int_{\partial E \cap \partial_D B} n \cdot \sigma \cdot (u - \bar{u}) + \int_{\partial E \cap \partial_N B} \bar{T} \cdot u. \end{aligned} \tag{2.13}$$

where u^{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, $I_h : V_h \times W_h^s \to \mathbb{R}$, is defined simply by summing over all elemental contributions,

$$I_h = \sum_{E \in \mathcal{T}_h} I_E. \tag{2.14}$$

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

$$\sum_{E\in\mathcal{T}_h}\int_E\Big(\delta\sigma\cdot\mathbb{C}^{-1}\cdot\sigma-\delta\sigma\cdot\nabla_s u\Big)+\int_\Gamma\{n\cdot\delta\sigma\}\cdot[\![u]\!]-\int_{\partial_D B}n\cdot\delta\sigma\cdot\bar{u}=0 \tag{2.15}$$

$$\sum_{E\in\mathcal{T}_h}\int_E\Big(-\sigma\cdot\nabla_s\delta u+f\cdot\delta u\Big)+\int_\Gamma\{n\cdot\sigma\}\cdot[\![\delta u]\!]+\int_{\partial_N B}\bar{T}\cdot\delta u=0. \tag{2.16}$$

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

$$\begin{aligned}\sum_{E\in\mathcal{T}_h}\int_E\Big(\gamma_h\cdot\mathbb{C}^{-1}\cdot\sigma_h-\gamma_h\cdot\nabla_s u_h\Big)+\int_\Gamma\{n\cdot\gamma_h\}\cdot[\![u_h]\!]&\\ =\int_{\partial_D B}n\cdot\gamma_h\cdot\bar{u},\qquad&\forall\gamma_h\in W_h^s;\end{aligned} \tag{2.17}$$

$$\begin{aligned}\sum_{E\in\mathcal{T}_h}\int_E\sigma_h\cdot\nabla_s v_h-\int_\Gamma\{n\cdot\sigma_h\}\cdot[\![v_h]\!]&\\ =\int_B f\cdot v_h+\int_{\partial_N B}\bar{T}\cdot v_h,\qquad&\forall v_h\in V_h.\end{aligned} \tag{2.18}$$

Next, we define the lifting operator $R_{\bar{u}}:(L^2(\Gamma))^d\to W_h^s$ by

$$\int_B R_{\bar{u}}(v)\cdot\gamma=-\int_\Gamma\{n\cdot\gamma\}\cdot v+\int_{\partial_D B}n\cdot\gamma\cdot\bar{u},\qquad\forall\gamma\in W_h^s. \tag{2.19}$$

This operator will now be used to derive the primal form [9] of the discretization where a single equation is obtained by eliminating σ_h between (2.17) and (2.18). In terms of (2.19), equation (2.17) is the same as

$$\sum_{E\in\mathcal{T}_h}\int_E\Big(\gamma_h\cdot\mathbb{C}^{-1}\cdot\sigma_h-\gamma_h\cdot\nabla_s u_h-R_{\bar{u}}([\![u_h]\!])\cdot\gamma_h\Big)=0\qquad\forall\gamma_h\in W_h^s.$$

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,

$$\sigma_h=\sigma_h(u_h)=\mathbb{C}\cdot\nabla_s u_h+\mathbb{C}\cdot R_{\bar{u}}([\![u_h]\!])\quad\text{ in } W_h^s. \tag{2.20}$$

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 (2.17) to get

$$\sum_{E\in\mathcal{T}_h}\int_E\Big(\nabla_s v_h\cdot\sigma_h-\nabla_s v_h\cdot\mathbb{C}\cdot\nabla_s u_h\Big)+\int_\Gamma\{n\cdot\mathbb{C}\cdot\nabla_s v_h\}\cdot[\![u_h]\!]=\int_{\partial_D B}n\cdot(\mathbb{C}\cdot\nabla_s v_h)\cdot\bar{u}.$$

Finally, substitute equation (2.18) to obtain

$$\begin{aligned}\sum_{E\in\mathcal{T}_h}\int_E\nabla_s v_h\cdot\mathbb{C}\cdot\nabla_s u_h-\int_\Gamma\Big(\{n\cdot\mathbb{C}\cdot\nabla_s v_h\}\cdot[\![u_h]\!]+\{n\cdot\sigma_h\}\cdot[\![v_h]\!]\Big)&\\ =\int_B f\cdot v_h+\int_{\partial_N B}\bar{T}\cdot v_h-\int_{\partial_D B}n\cdot(\mathbb{C}\cdot\nabla_s v_h)\cdot\bar{u}\,.&\end{aligned} \tag{2.21}$$

If $(u_h, \sigma_h) \in V_h \times W_h^s$ solves (2.17) - (2.18), then u_h solves (2.21), with $\sigma_h = \sigma_h(u_h)$ given by (2.20). Equation (2.21) is called the primal formulation.

Recall the definition of $R_{\bar{u}}$, (2.19), and introduce the notation $R = R_0$. Using (2.19) and (2.20), the primal form (2.21) can also be written as

$$\sum_{E\in\mathcal{T}_h} \int_E (\nabla_s v_h + R([\![v_h]\!])) \cdot \mathbb{C} \cdot (\nabla_s u_h + R([\![u_h]\!])) =$$
$$\int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h - \int_{\partial_D B} n \cdot \Big(\mathbb{C}\cdot(\nabla_s v_h + R([\![v_h]\!]))\Big) \cdot \bar{u}. \tag{2.22}$$

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 [15, 17]. Arnold, *et al.* [9] show that this discretization is consistent, conservative and adjoint consistent, but unstable, for the scalar Poisson equation, and we show in Section 3 that these properties carry over to linear elasticity.

Brezzi, *et al.*, [16, 17] propose a stabilizing term for the scalar case which naturally extends to linear elasticity. The stabilization is given in terms of $r_{e,\bar{u}} : (L^2(\Gamma))^d \to W_h^s$. Define $r_{e,\bar{u}}$ for $e \in \mathcal{E}_h^I$,

$$\int_B r_{e,\bar{u}}(v) \cdot \gamma = -\int_e \{n \cdot \gamma\} \cdot v, \qquad \forall \gamma \in W_h^s, \tag{2.23}$$

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

$$\int_B r_{e,\bar{u}}(v) \cdot \gamma = -\int_e \{n \cdot \gamma\} \cdot v + \int_e n \cdot \gamma \cdot \bar{u}, \qquad \forall \gamma \in W_h^s, \tag{2.24}$$

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 $E \in \mathcal{T}_h$,

$$R_{\bar{u}}(v) = \sum_{e\subset\partial E} r_{e,\bar{u}}(v) \tag{2.25}$$

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

$$\sum_{E\in\mathcal{T}_h} \int_E (\nabla_s v_h + R([\![v_h]\!])) \cdot \mathbb{C} \cdot (\nabla_s u_h + R([\![u_h]\!])) + \beta \sum_{e\in\mathcal{E}_h} \int_B r_e([\![u_h]\!]) \cdot \mathbb{C} \cdot r_e([\![v_h]\!])$$
$$= \int_B f \cdot v_h + \int_{\partial_N B} \bar{T} \cdot v_h - \int_{\partial_D B} n \cdot \Big(\mathbb{C} \cdot (\nabla_s v_h + R([\![v_h]\!]) + \beta r_e([\![v_h]\!]))\Big) \cdot \bar{u}. \tag{2.26}$$

The form (2.26), 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, *et al.*, [16, 17], 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, ∂B. Thus, $\partial_N B = \emptyset$,

$\mathcal{E}_h^N = \emptyset$, and without loss of generality, $\bar{u} = 0$ on ∂B. Accordingly, the complete discrete problem statement, with these modifications, is to find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = \int_B f \cdot v_h \qquad \forall v_h \in V_h \tag{2.27}$$

where the bilinear form a_h is given by

$$\begin{aligned} a_h(u_h, v_h) = \sum_{E \in \mathcal{T}_h} \int_E \Big(\nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u_h + \nabla_s v_h \cdot \mathbb{C} \cdot R([\![u_h]\!]) + R([\![v_h]\!]) \cdot \mathbb{C} \cdot \nabla_s u_h \Big) \\ + \beta \sum_{e \in \mathcal{E}_h} \int_B r_e([\![u_h]\!]) \cdot \mathbb{C} \cdot r_e([\![v_h]\!]). \end{aligned} \tag{2.28}$$

Remark. Both problems, (2.26) and (2.28), can be written in a two-field form, *i.e.*, with both u_h and σ_h as unknowns.

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 $H^p(\Omega)$ semi-norm, 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 $\hat{V} = (H_0^1(B))^d + V_h$ given by

$$|\!|\!| v |\!|\!|_s^2 = \sum_{E \in \mathcal{T}_h} \|\nabla_s v\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e([\![v]\!])\|_{0,B}^2 \qquad v \in \hat{V} \tag{2.29}$$

$$|\!|\!| v |\!|\!|^2 = \sum_{E \in \mathcal{T}_h} \|\nabla v\|_{0,E}^2 + \sum_{e \in \mathcal{E}_h} \|r_e([\![v]\!])\|_{0,B}^2 \qquad v \in \hat{V}. \tag{2.30}$$

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

$$|\!|\!| v |\!|\!|_s^2 \le |\!|\!| v |\!|\!|^2 \qquad v \in \hat{V} \tag{2.31}$$

which shows that $|\!|\!| \cdot |\!|\!|$ is also a norm on $\hat{V}$. Although one might expect the second term in the definition of norms (2.29) and (2.30) to act as an L^2-like contribution, we can only assert that these are semi-norms on $\tilde{V}$. In the case of the scalar Poisson equation, [9, 12, 13, 14, 16, 17, 18], there is no need to distinguish between the norms (2.29) and (2.30). Following the ideas in [16, 17], 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 $L^2(B)$ (Theorem 3.12 and Theorem 3.13). The convergence in the $|\!|\!| \cdot |\!|\!|_s$-norm is sufficient for a mesh-independent BD estimate (Theorem 3.16); 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, [23].

In order to obtain convergence in the norm, $|\!|\!| \cdot |\!|\!|$, Theorem 3.20, we prove a generalized version of Korn's second inequality for the subdivision, Corollary 3.19.

The proof of this inequality relies on observations about how Korn's inequality for an element behaves under distortion (Theorem 3.17) and scaling (Theorem 3.18). Finally, Theorem 3.22 shows that the mesh-dependent norm, $\|\cdot\|$, estimates the BV norm, and as a consequence, Corollary 3.23, we obtain convergence in BV, an optimal mesh-independent result.

3. Theoretical Results.

3.1. Convergence in the mesh-dependent symmetric norm. Following the developments in [16, 17] for the two-dimensional Poisson equation, we obtain the convergence of the discretized solutions in the mesh-dependent norm $\|\cdot\|_s$. 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 ϕ 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 $\phi \in T(e)$,*

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

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

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

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 $\|\gamma_h\|^2_{0,\hat{E}}$ subject to the linear constraint $\gamma_h|_e = \phi \in T(\hat{e})$, which depends continuously on ϕ. 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

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

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 $\|F\| \le h/\hat{\rho}$ (see e.g. [21, p. 120]), where $\hat{\rho}$ is the diameter of the largest ball contained in $\hat{E}$, to obtain

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

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*

$$\|r_e(v)\|_{0,e} \leq C h^{-1/2} \|r_e(v)\|_{0,E}, \tag{3.4}$$

for all $v \in (L^2(e))^d$.

Proof. The inequality (3.4) 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

$$\|\hat{\gamma}\|_{0,\hat{e}} \leq C\ \|\hat{\gamma}\|_{0,\hat{E}} \tag{3.5}$$

for all $\hat{\gamma} \in (W_h^{\hat{E}})^{d\times d}$. Inequality (3.5) is a direct consequence of the continuity of the trace in $W_h^{\hat{E}} \subset H^1(\hat{E})$ (see, e. g., [22, pag. 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,

$$\|\gamma\|_{0,E}^2 = \int_E \gamma\cdot\gamma\ = |\det F|\ \int_{\hat{E}} \hat{\gamma}\cdot\hat{\gamma}\ = |\det F|\ \|\hat{\gamma}\|_{0,\hat{E}}^2 \tag{3.6}$$

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

where $\hat{n}$ is the unit outward normal to $\hat{e}$. Therefore, (3.5), (3.6), and (3.7) combine to yield

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

The last part of the bound uses the fact that $\|F^{-1}\| \leq \hat{h}/(\mathrm{diam}\ B_E) \leq \hat{h}/(\rho h)$ (see e.g. [21, p. 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*

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

and

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

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, it is possible to choose γ_h^e so that $\|\gamma_h^e\| \leq C\|[\![v_h]\!]\|$. For the tensor γ_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 (2.23) to get

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

In the nontrivial case in which $\|[\![v_h]\!]\|_{0,e} \neq 0$, the 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 (2.23) to get

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

We have used the linearity of r_e and (3.4) in the last step. The result, (3.10), follows. ☐

Proposition 3.4 (Symmetric norm). *Let $v_h \in \hat{V} = (H^1_0(B))^d + V_h$. Then $\|\cdot\|_s : \hat{V} \to \mathbb{R}$ as defined in (2.29) is a norm on $\hat{V}$.*

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 $\hat{V}$. Notice that $\|v\|_s = 0$ iff $\|\nabla_s v\|_{0,E} = 0$ for all $E \in \mathcal{T}_h$ and $\|r_e([\![v]\!])\|_{0,B} = 0$ for all $e \in \mathcal{E}_h$. Let $v = v_1 + v_2 \in \hat{V}$, with $v_1 \in (H^1_0(B))^d$ and $v_2 \in V_h$. By Lemma 3.3, we have that $\|[\![v_2]\!]\|_{0,e} \leq C\, h^{1/2} \|r_e([\![v_2]\!])\|_{0,B}$. Therefore, $\|[\![v_2]\!]\|_{0,e} = 0$. Since also $\|[\![v_1]\!]\|_{0,e} = 0$ we have that $\|[\![v]\!]\|_{0,e} = 0$. So $v \in (H^1_0(B))^d$ by [24, Theorem 1.3]. Korn's first inequality for homogeneous boundary data applied to $v \in (H^1_0(B))^d$ then shows that $v = 0$. ☐

Next, we show that the bilinear form (2.28) is continuous and coercive with respect to the norm, $\|\cdot\|_s$. The proofs follow [16, 17] 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*

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

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

$$\text{(ii)} \quad a_h(u_h, u_h) \geq \mu \, \|u_h\|^2_s \qquad \forall\, u_h \in \hat{V}.$$

Proof. We first prove the following inequality, a consequence of equation (2.25),

$$\|R([\![v_h]\!])\|^2_{0,E} \leq N_e \sum_{e \subset \partial E} \|r_e([\![v_h]\!])\|^2_{0,E}. \tag{3.12}$$

We have

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

Next, the continuity of the bilinear form (2.28) follows from estimating each term.

$$\left| \int_E \nabla_s u_h \cdot \mathbb{C} \cdot \nabla_s v_h \right| \leq \|\mathbb{C}\| \, \|\nabla_s u_h\|_{0,E} \, \|\nabla_s v_h\|_{0,E} \tag{3.13}$$

$$\begin{aligned} \left| \int_E \nabla_s u_h \cdot \mathbb{C} \cdot R([\![v_h]\!]) \right| &\leq \|\mathbb{C}\| \, \|\nabla_s u_h\|_{0,E} \, \|R([\![v_h]\!])\|_{0,E} \\ &\leq \|\mathbb{C}\| \, \|\nabla_s u_h\|_{0,E} \left[N_e \sum_{e \subset \partial E} \|r_e([\![v_h]\!])\|^2_{0,E} \right]^{1/2} \end{aligned} \tag{3.14}$$

$$\left| \sum_{e \subset \partial E} \int_E r_e([\![u_h]\!]) \cdot \mathbb{C} \cdot r_e([\![v_h]\!]) \right| \leq \|\mathbb{C}\| \sum_{e \subset \partial E} \|r_e([\![u_h]\!])\|_{0,E} \, \|r_e([\![v_h]\!])\|_{0,E} \, . \tag{3.15}$$

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

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

$$\|\gamma\|^2_{0,E,\mathbb{C}} = \int_E \gamma \cdot \mathbb{C} \cdot \gamma \qquad \forall \gamma \in W^s_h.$$

Due to (3.12), we get

$$\begin{aligned} a_h(u_h, u_h) &= \sum_{E \in \mathcal{T}_h} \left(\|\nabla_s u_h\|^2_{0,E,\mathbb{C}} + \int_E 2 \nabla_s u_h \cdot \mathbb{C} \cdot R([\![u_h]\!]) + \beta \sum_{e \subset \partial E} \|r_e([\![u_h]\!])\|^2_{0,E,\mathbb{C}} \right) \\ &\geq \sum_{E \in \mathcal{T}_h} \left((1-\varepsilon) \|\nabla_s u_h\|^2_{0,E,\mathbb{C}} - \frac{1}{\varepsilon} \|R([\![u_h]\!])\|^2_{0,E,\mathbb{C}} + \beta \sum_{e \subset \partial E} \|r_e([\![u_h]\!])\|^2_{0,E,\mathbb{C}} \right) \\ &\geq \sum_{E \in \mathcal{T}_h} \left((1-\varepsilon) \|\nabla_s u_h\|^2_{0,E,\mathbb{C}} + \left(\beta - \frac{N_e}{\varepsilon} \right) \sum_{e \subset \partial E} \|r_e([\![u_h]\!])\|^2_{0,E,\mathbb{C}} \right), \end{aligned}$$

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

Remark. As suggested in [16, 17], following the same steps as in the previous proof establishes continuity and coercivity of the bilinear form given by equation (2.26), but for any $\beta > 0$.

In addition to being continuous and coercive, the bilinear form (2.28) is consistent and adjoint consistent (as is (2.26)). Consistency is the requirement that the exact solution of the partial differential equation be a solution of the discrete problem, and similarly for adjoint consistency [9]. A precise definition is given in the propositions below. These properties form the basis for establishing convergence of the discrete displacement, first in $\|\!|\cdot|\!\|_s$, and subsequently in $L^2(B)$. The following lemma is a preliminary to proving consistency.

LEMMA 3.6. *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*

$$\sum_{E \in \mathcal{T}_h} \int_{\partial E} n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h = \sum_{e \in \mathcal{E}_h} \int_e \{ n \cdot \mathbb{C} \cdot \nabla u \} \cdot [\![v_h]\!] \, . \tag{3.16}$$

Proof. The assumed regularity of u implies that $n \cdot \mathbb{C} \cdot \nabla u$ is continuous across inter-element boundaries (e.g., [24, Theorem 1.3]); *i.e.*, $0 = n \cdot (\mathbb{C} \cdot \nabla u^- - \mathbb{C} \cdot \nabla u^+)$ on any face in $\mathcal{E}_h^I$. Therefore

$$
\begin{aligned}
\sum_{E \in \mathcal{T}_h} \int_{\partial E} & n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\
&= \sum_{e \in \mathcal{E}_h^I} \int_e \Big(- n \cdot (\mathbb{C} \cdot \nabla u^+) \cdot v_h^+ + n \cdot (\mathbb{C} \cdot \nabla u^-) \cdot v_h^- \Big) + \sum_{e \in \mathcal{E}_h^D} \int_e n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\
&= \sum_{e \in \mathcal{E}_h^I} \int_e \Big(- \frac{1}{2} \Big(n \cdot (\mathbb{C} \cdot \nabla u^+) + n \cdot (\mathbb{C} \cdot \nabla u^-) \Big) \cdot v_h^+ + \frac{1}{2} \Big(n \cdot (\mathbb{C} \cdot \nabla u^+) + n \cdot (\mathbb{C} \cdot \nabla u^-) \Big) \cdot v_h^- \\
&\qquad + \sum_{e \in \mathcal{E}_h^D} \int_e n \cdot (\mathbb{C} \cdot \nabla u) \cdot v_h \\
&= \sum_{e \in \mathcal{E}_h} \int_e \{ n \cdot \mathbb{C} \cdot \nabla u \} \cdot [\![v_h]\!] .
\end{aligned}
$$

□

PROPOSITION 3.7 (Consistency of the bilinear form). *Let u be the exact solution of (2.1), with $u \in (H^m(B))^d$ for some m such that $2 \leq m \leq k+1$, then*

$$
a_h(u, v_h) = \int_B f \cdot v_h \quad \forall v_h \in V_h, \tag{3.17}
$$

where the bilinear form is given in (2.28).

Proof. To establish consistency of the bilinear form, multiply (2.1) by $v_h \in V_h$ and integrate by parts over each element,

$$
\begin{aligned}
\int_B f \cdot v_h &= - \int_B v_h \cdot \nabla \cdot (\mathbb{C} \cdot \nabla_s u) \\
&= \sum_{E \in \mathcal{T}_h} \Big(\int_E \nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u - \int_{\partial E} n \cdot (\mathbb{C} \cdot \nabla_s u) \cdot v_h \Big) \\
&= \sum_{E \in \mathcal{T}_h} \int_E \nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u - \sum_{e \in \mathcal{E}_h} \int_e \{ n \cdot \mathbb{C} \cdot \nabla_s u \} \cdot [\![v_h]\!] - \sum_{e \in \mathcal{E}_h} \{ n \cdot \mathbb{C} \cdot \nabla_s v_h \} \cdot [\![u]\!] \\
&= \sum_{E \in \mathcal{T}_h} \int_E \Big(\nabla_s v_h \cdot \mathbb{C} \cdot \nabla_s u + R([\![v_h]\!]) \cdot \mathbb{C} \cdot \nabla_s u + \nabla_s v_h \cdot \mathbb{C} \cdot R([\![u]\!]) \Big) \\
&\qquad + \beta \sum_{e \in \mathcal{E}_h} \int_B r_e([\![u]\!]) \cdot \mathbb{C} \cdot r_e([\![v_h]\!]) \\
&= a_h(u, v_h).
\end{aligned}
$$

We have used Lemma 3.6, the fact that $[\![u]\!] = 0$, and (2.19). □

COROLLARY 3.8 (Galerkin orthogonality). *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 $u_h \in V_h$ solve (2.27), then*

$$
a_h(u - u_h, v_h) = 0, \qquad \forall\, v_h \in V_h. \tag{3.18}
$$

Proof. Subtract the consistency condition, (3.17), and the characterization of the discrete solution, (2.27) to establish the desired result. □

Since the problem (2.1) is self-adjoint, the adjoint problem is the same as the original problem; namely, find $w \in (H^2(B))^d$ such that

$$\begin{cases} -\nabla\cdot(\mathbb{C}\cdot\nabla_s w) = g & \text{in } B \\ w = 0 & \text{on } \partial B \end{cases} \tag{3.19}$$

for $g \in (L^2(B))^d$.

COROLLARY 3.9 (Adjoint consistency). *Let $w \in (H^2(B))^d$ be the solution to the adjoint problem (3.19), then*

$$a_h(v_h, w) = \int_B g\cdot v_h \quad \forall v_h \in V_h. \tag{3.20}$$

Proof. Since the problem is self-adjoint, the condition (3.20) follows from consistency. □

The condition (3.20) on the bilinear form is called adjoint consistency [9].

The last component required to prove convergence is a bound on the approximation error $|||u-u_I|||_s$ when u_I is a suitable interpolant of the exact solution u. Arnold, *et al.*, [9], note that discontinuous interpolants can be employed, if they satisfy a local approximation property summarized in the next theorem.

THEOREM 3.10 (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*

$$|v - v_I|_{q,E} \le Ch^{k+1-q}|v|_{k+1,E}, \quad k+1 \ge q \ge 0, \tag{3.21}$$

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

Proof. Ciarlet, [21, Theorem 3.1.5]. □

THEOREM 3.11 (Interpolation-error estimate). *Let $u \in (H^m(B))^d$ for some m such that $2 \le m \le k+1$, 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,*

$$|||u-u_I|||_s \le Ch^{m-1}|u|_{m,B}, \tag{3.22}$$

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.

Proof. From the previous theorem, we have

$$\sum_{E\in\mathcal{T}_h} |u-u_I|^2_{q,E} \le \sum_{E\in\mathcal{T}_h} Ch^{2m-2q}|u|^2_{m,E}, \quad m \ge q. \tag{3.23}$$

In addition, the trace inequality [25, p. 133] together with a scaling argument gives

$$\|u\|^2_{0,e} \le C\Big(h^{-1}|u|^2_{0,E} + h|u|^2_{1,E}\Big) \qquad \forall u \in H^1(E) \tag{3.24}$$

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 [9], the interpolation inequality (3.22) is established using the inequality (3.24), the bound (3.23), and the inverse inequality (3.10). Starting from the definition of $\|\!|\cdot\|\!|_s$, the theorem is obtained as follows,

$$
\begin{aligned}
\|\!|u-u_I\|\!|_s^2 &= \sum_{E\in\mathcal{T}_h} \|\nabla_s(u-u_I)\|_{0,E}^2 + \sum_{e\in\mathcal{E}_h} \|r_e([\![u-u_I]\!])\|_{0,B}^2 \\
&\le \sum_{E\in\mathcal{T}_h} \|\nabla(u-u_I)\|_{0,E}^2 + \sum_{e\in\mathcal{E}_h} \|r_e([\![u-u_I]\!])\|_{0,B}^2 \\
&\le \sum_{E\in\mathcal{T}_h} |u-u_I|_{1,E}^2 + \sum_{e\in\mathcal{E}_h} Ch^{-1}\|[\![u-u_I]\!]\|_{0,e}^2 \\
&\le C\sum_{E\in\mathcal{T}_h} h^{2m-2}|u|_{m,E}^2 \\
&\le Ch^{2m-2}|u|_{m,B}^2. \qquad (3.25)
\end{aligned}
$$

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. □

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 3.12 (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 $2\le m\le k+1$, and let u_h be the solution of (2.27), then the following estimate holds*

$$\|\!|u-u_h\|\!|_s \le C\,h^{m-1}\,|u|_{m,B}, \qquad (3.26)$$

where C is a positive constant independent of h.

Proof. From Proposition 3.5, we have

$$
\begin{aligned}
\mu\|\!|u_I-u_h\|\!|_s^2 &\le a_h(u_I-u_h,u_I-u_h) \\
&= a_h(u_I-u,u_I-u_h)+a_h(u-u_h,u_I-u_h) \\
&\le M\|\!|u_I-u_h\|\!|_s\|\!|u_I-u\|\!|_s + a_h(u-u_h,u_I-u_h) \\
&= M\|\!|u_I-u_h\|\!|_s\|\!|u_I-u\|\!|_s. \qquad (3.27)
\end{aligned}
$$

Note, $a_h(u-u_h,u_I-u_h)=0$ follows from Galerkin orthogonality, (3.18). Insert (3.22) into (3.27) to obtain the desired result. □

THEOREM 3.13 (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 $2\le m\le k+1$, and let u_h be the solution of (2.27), then the following estimate holds*

$$\|u-u_h\|_{0,B}\le C\;h^m|u|_{m,B}\,. \qquad (3.28)$$

Proof. The proof follows a standard duality argument. Consider the adjoint problem (3.19), with $g=u-u_h$. Now, take $w_I\in V_h$ to be the piecewise linear interpolant of w over each element, and use $v_h=u-u_h$ in (3.20) to obtain,

$$
\begin{aligned}
\|u-u_h\|_{0,B}^2 &= a_h(u-u_h,w) \\
&= a_h(u-u_h,w-w_I) \\
&\le M\|\!|u-u_h\|\!|_s\|\!|w-w_I\|\!|_s \\
&\le Ch|w|_{2,B}\|\!|u-u_h\|\!|_s,
\end{aligned}
$$

where we have used (3.18) and Theorem 3.11 for the interpolation error estimate $\|w - w_I\|_s$. Since $u - u_h \in (L_2(B))^d$, the following standard elliptic regularity estimate holds (see, e.g., [26]),

$$\|w\|_{2,B} \leq C\|u - u_h\|_{0,B} \tag{3.29}$$

for some constant $C > 0$, and the theorem follows. □

COROLLARY 3.14 (Convergence of the stress in $L^2(B)$). *Let σ be the exact solution with components in $H^{m-1}(B)$ for some m such that $2 \leq m \leq k+1$, and let σ_h be given by (2.20), then the following estimate holds*

$$\|\sigma - \sigma_h\|_{0,B} \leq Ch^{m-1}|u|_{m,B}. \tag{3.30}$$

Proof. For the exact solution $[\![u]\!] = 0$, which implies $R([\![u]\!]) = 0$. So we can write $\sigma = \mathbb{C} \cdot (\nabla_s u + R([\![u]\!]))$. Therefore,

$$\sigma - \sigma_h = \mathbb{C} \cdot \nabla_s(u - u_h) + \mathbb{C} \cdot R([\![u - u_h]\!]). \tag{3.31}$$

It follows that

$$\begin{aligned}
\|\sigma - \sigma_h\|_{0,B}^2 &= \sum_{E \in \mathcal{T}_h} \|\sigma - \sigma_h\|_{0,E}^2 \\
&= \sum_{E \in \mathcal{T}_h} \|\mathbb{C} \cdot \nabla_s(u - u_h) + \mathbb{C} \cdot R([\![u - u_h]\!])\|_{0,E}^2 \\
&\leq \sum_{E \in \mathcal{T}_h} C\big(\|\nabla_s(u - u_h)\|_{0,E}^2 + N_e \sum_{e \subset \partial E} \|r_e([\![u - u_h]\!])\|_{0,E}^2\big) \\
&\leq C\|u - u_h\|_s^2 \leq Ch^{2m-2}|u|_{m,B}^2.
\end{aligned}$$

□

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 (2.20), and is not, in general, proportional to the strain.

Remark. Again, as suggested in [16, 17], it can also be proved that the bilinear form given by equation (2.26) is both consistent and adjoint consistent. Therefore, the same error estimates hold for the problem directly derived from the variational principle, equation (2.26).

3.2. The natural (suboptimal but mesh-independent) BD-estimate. Possible discontinuities in the displacement across element boundaries naturally leads 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, $\mathcal{E}(Du) = \frac{1}{2}(Du + Du^T)$, 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 3.15 (Poincaré for BD). *Let $B \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary. Then*

$$\exists\, C > 0: \quad \forall\, u \in BD(B),\ u_{|\partial B} = 0, \quad \|u\|_{L^1(B)} \leq C\, \|\mathcal{E}(Du)\|(B), \tag{3.32}$$

where $u_{|\partial B}$ denotes the generalized trace.

Proof. Témam [27, p. 189], Remark II.2.5. ☐

THEOREM 3.16 (Natural BD estimate).

$$\exists\, C > 0 : \quad \forall\, u \in \hat{V}, \quad \|u\|_{BD(B)} \leq C \,|\!|\!| u |\!|\!|_s, \tag{3.33}$$

with C independent of h.

Proof. Recall the definition of the BD norm

$$\|u\|_{BD(B)} = \|u\|_{L^1(B)} + \|\mathcal{E}(Du)\|(B)$$

where

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

The proof continues *mutatis mutandis* as in Theorem 3.22 below. ☐

Using the estimate (3.33) for the difference $u - u_h$ together with Theorem 3.11 shows that convergence of the method is immediately strengthened from the $|\!|\!| \cdot |\!|\!|_s$-norm to a mesh-independent estimate in the space $BD(B)$. It is clear that any 'optimal' estimate in the symmetric norm, derived under less smoothness assumptions on the underlying continuous problem [20], 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 3.13 that additionally establishes convergence of the discrete solutions in $L^2(B)$.

The occurrence of the space BD is, strictly speaking, an artifact of the linearized treatment where only the symmetrized infinitesimal strains $\varepsilon(\nabla u)$ appear. Since this BD estimate does not control the antisymmetric part of the displacement gradient, we are interested in obtaining convergence in the space $BV(B)$. However, since $BV(B)$ is strictly smaller than $BD(B)$ there is no obvious way to proceed directly from the BD estimate to a BV estimate. Instead, we will first strengthen Theorem 3.12 to the $|\!|\!| \cdot |\!|\!|$-norm. Note that for a given mesh size $h > 0$, given the finite dimensionality of V_h and the fact that both $|\!|\!| \cdot |\!|\!|$ and $|\!|\!| \cdot |\!|\!|_s$ are norms in V_h, we have for $u_h \in V_h$,

$$\|u_h\|_{BD} \leq \|u_h\|_{BV} \leq C |\!|\!| u_h |\!|\!| \leq c(h) |\!|\!| u_h |\!|\!|_s, \tag{3.34}$$

where the estimate $\|u_h\|_{BV} \leq C |\!|\!| u_h |\!|\!|$ is obtained in Theorem 3.22. However, $c(h)$ may not be bounded from below away from zero for all $h > 0$, the possibility of which has been observed numerically. 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 [23]. 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 $\mathrm{SL}(d, \mathbb{R}) = \{X \in \mathbb{R}^{d\times d} \mid \det X = 1\}$.

THEOREM 3.17 (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 $\mathcal{M} = \{X \in \mathrm{SL}(d, \mathbb{R}) : \|X\| \leq K\}$, for some $K > 0$. For $F \in \mathcal{M}$ define $\Omega_\xi = F(\Omega)$. Then*

$$\exists\, C > 0 : \quad \forall\, F \in \mathcal{M}, \quad \forall\, u \in H^1(\Omega_\xi),$$
$$\|\nabla_\xi u^{\mathsf{T}} + \nabla_\xi u\|^2_{0,\Omega_\xi} + \|u\|^2_{0,\Omega_\xi} \geq C\, \|u\|^2_{1,\Omega_\xi}.$$

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

$$(3.35)\qquad \int_{\Omega_\xi} \|\nabla_\xi u^\mathsf{T} + \nabla_\xi u\|^2 + \|u\|^2 = \int_\Omega \|F^{-\mathsf{T}}\nabla\tilde{u}^\mathsf{T} + \nabla\tilde{u}\, F^{-1}\|^2 + \|\tilde{u}\|^2 \,.$$

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

$$(3.36)\qquad \|F_n^{-\mathsf{T}}\nabla\tilde{u}_n^\mathsf{T} + \nabla\tilde{u}_n\, F_n^{-1}\|_{0,\Omega}^2 + \|\tilde{u}_n\|_{0,\Omega}^2 \le \frac{1}{n}\,\|\tilde{u}_n\|_{1,\Omega}^2 = \frac{1}{n}.$$

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

$$(3.37)\qquad \|\hat{F}^{-\mathsf{T}}\nabla\tilde{u}_n^\mathsf{T} + \nabla\tilde{u}_n\hat{F}^{-1}\|_{0,\Omega}^2 + \|\tilde{u}_n\|_{0,\Omega}^2 \to 0.$$

Thus $\tilde{u}_n$ is a minimizing sequence. For fixed $\hat{F}$ the quadratic expression is uniformly positive (generalized Korn's second inequality, see [28]) such that

$$(3.38)\qquad \|\hat{F}^{-\mathsf{T}}\nabla\tilde{u}_n^\mathsf{T} + \nabla\tilde{u}_n\hat{F}^{-1}\|_{0,\Omega}^2 + \|\tilde{u}_n\|_{0,\Omega}^2 \ge C(\hat{F})\,\|\tilde{u}_n\|_{1,\Omega}^2$$

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

THEOREM 3.18 (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*

$$\exists\, C(\Omega) > 0 : \quad \forall\, u \in H^1(\Omega_h),$$
$$\|\nabla u^\mathsf{T} + \nabla u\|_{0,\Omega_h}^2 + \frac{1}{|\Omega_h|^{2/d}}\,\|u\|_{0,\Omega_h}^2 \ge C(\Omega)\left(\|\nabla u\|_{0,\Omega_h}^2 + \frac{1}{|\Omega_h|^{2/d}}\,\|u\|_{0,\Omega_h}^2\right),$$

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

Proof. Let $\tilde{u} \in H^1(\Omega)$. From Korn's second inequality (see, e.g.,[28]) we get

$$(3.39)\qquad \|\nabla\tilde{u}^\mathsf{T} + \nabla\tilde{u}\|_{0,\Omega}^2 + \|\tilde{u}\|_{0,\Omega}^2 \ge C(\Omega)\left(\|\nabla\tilde{u}\|_{0,\Omega}^2 + \|\tilde{u}\|_{0,\Omega}^2\right).$$

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

$$\frac{1}{h^{d-2}}\|\nabla u^\mathsf{T} + \nabla u\|_{0,\Omega_h}^2 + \frac{1}{h^d}\,\|u\|_{0,\Omega_h}^2 \ge C(\Omega)\left(\frac{1}{h^{d-2}}\|\nabla u\|_{0,\Omega_h}^2 + \frac{1}{h^d}\|u\|_{0,\Omega_h}^2\right)$$

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

COROLLARY 3.19 (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,*

$$\exists C > 0 : \quad \forall\, E \in (\mathcal{T}_h), \quad \forall\, u \in H^1(E),$$
$$\|\nabla u^\mathsf{T} + \nabla u\|_{0,E}^2 + \frac{1}{|E|^{2/d}}\,\|u\|_{0,E}^2 \ge C\left(\|\nabla u\|_{0,E}^2 + \frac{1}{|E|^{2/d}}\,\|u\|_{0,E}^2\right).$$

Proof. Let F be an affine transformation such that $E = F(\hat{E})$. 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 Theorem 3.1.3 in [29, p. 120] and the quasi-uniformity of the subdivision we have that

$$\|\tilde{F}\| = \frac{\|F\|}{(\det F)^{1/d}} \leq \frac{h}{\hat{\rho}} \frac{1}{|E|^{1/d}} \leq \frac{C}{\hat{\rho}} \tag{3.40}$$

where $\hat{\rho}$ is the diameter of the largest ball contained in $\hat{E}$ and C is independent of E. Therefore, by Theorem 3.17 we can state Korn's second inequality for each domain $\tilde{F}(\hat{E})$ in the subdivision with the same constant $C > 0$. The corollary then follows from Theorem 3.18. ◻

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 3.20 (Convergence in the mesh-dependent norm $\|\!|\cdot|\!\|$). *Let $(v_h) \subset V_h$ be a sequence such that $\|\!|v_h|\!\|_s \leq C h^{m-1}$ and $\|v_h\|_{0,B} \leq C h^m$ for $h \downarrow 0$. Then*

$$\|\!|v_h|\!\| \leq C\, h^{m-1} \tag{3.41}$$

for some $C > 0$ independent of h.

Proof. Use Corollary 3.19 and sum over the elements to obtain the estimate

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

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

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

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

$$\begin{aligned} &\sum_{E\in\mathcal{T}_h} \|\nabla v_h^{\mathsf{T}} + \nabla v_h\|_{0,E}^2 + \frac{1}{h^2}\, \|v_h\|_{0,E}^2 + \sum_{e\in\mathcal{E}_h} \|r_e([\![v_h]\!])\|_{0,B}^2 \\ &\qquad \geq C \sum_{E\in\mathcal{T}_h} \left(\|\nabla v_h\|_{0,E}^2 + \frac{1}{h^2}\, \|v_h\|_{0,E}^2 + \sum_{e\in\mathcal{E}_h} \|r_e([\![v_h]\!])\|_{0,B}^2 \right) \end{aligned} \tag{3.42}$$

or

$$\|\!|v_h|\!\|_s^2 + \frac{1}{h^2} \sum_{E\in\mathcal{T}_h} \|v_h\|_{0,E}^2 \geq C \left(\|\!|v_h|\!\|^2 + \frac{1}{h^2} \sum_{E\in\mathcal{T}_h} \|v_h\|_{0,E}^2 \right) \tag{3.43}$$

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

$$\|\!|v_h|\!\|_s^2 + \frac{1}{h^2}\, \|v_h\|_{0,B}^2 \geq C \left(\|\!|v_h|\!\|^2 + \frac{1}{h^2}\, \|v_h\|_{0,B}^2 \right) \geq C\, \|\!|v_h|\!\|^2. \tag{3.44}$$

Using the convergence of (v_h) and equation (3.44) we obtain

$$\|\!|v_h|\!\|^2 \leq C \left(h^{2m-2} + \frac{1}{h^2}\, h^{2m} \right) = C h^{2m-2} \tag{3.45}$$

which completes the theorem. ◻

Remark. As it is evident from the statement of Theorem 3.20, 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. 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 $\hat{V} = V_h + (H_0^1(B))^d$ and as a result, 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 3.21 (Poincaré for BV).

$$\exists\, C > 0: \quad \forall\, u \in BV(\mathbb{R}^d), \quad \|u\|_{L^{d/(d-1)}(\mathbb{R}^d)} \le C\, \|Du\|(\mathbb{R}^d). \tag{3.46}$$

Proof. Evans and Gariepy, [25, p. 189] Theorem 1. □

THEOREM 3.22 (Natural BV estimate).

$$\exists\, C > 0: \quad \forall\, u \in \hat{V}, \quad \|u\|_{BV} \le C\, \|\!|u|\!\|,$$

with C independent of h.

Proof. Recall the definition of the BV norm

$$\|u\|_{BV(B)} = \|u\|_{L^1(B)} + \|Du\|(B) \tag{3.47}$$

where

$$\|Du\|(B) = \sup\left\{ \int_B u\cdot(\nabla\cdot\Psi) \;:\; \Psi \in C_0^1(B, \mathbb{R}^{d\times d}),\ \|\Psi\|_{L^\infty} \le 1 \right\}. \tag{3.48}$$

First observe that

$$\begin{aligned}\int_B u\cdot(\nabla\cdot\Psi) &= \sum_{E\in\mathcal{T}_h}\int_E u\cdot(\nabla\cdot\Psi) = \sum_{E\in\mathcal{T}_h}\int_E \nabla\cdot(\Psi\cdot u) - \sum_{E\in\mathcal{T}_h}\int_E \Psi\cdot\nabla u\\ &= \sum_{E\in\mathcal{T}_h}\int_{\partial E} n_E\cdot\Psi\cdot u - \sum_{E\in\mathcal{T}_h}\int_E \Psi\cdot\nabla u\\ &= \sum_{e\in\mathcal{E}_h}\int_e n\cdot\Psi\cdot[\![u]\!] - \sum_{E\in\mathcal{T}_h}\int_E \Psi\cdot\nabla u\end{aligned}$$

Each term in the two sums may be estimated individually by

$$\sup_{\|\Psi\|_{L^\infty}\le 1}\left[\int_e n\cdot\Psi\cdot[\![u]\!]\right] \le \int_e [\![u]\!]\cdot\frac{[\![u]\!]}{\|[\![u]\!]\|} \le \|[\![u]\!]\|_{L^1(e)}, \tag{3.49}$$

and

$$\sup_{\|\Psi\|_{L^\infty}\le 1}\left[-\int_E \Psi\cdot\nabla u\right] \le \int_E \frac{\nabla u}{\|\nabla u\|}\cdot\nabla u \le \|\nabla u\|_{L^1(E)}, \tag{3.50}$$

which yields the preliminary estimate

$$\|Du\|(B) \le \sum_{e\in\mathcal{E}_h}\|[\![u]\!]\|_{L^1(e)} + \sum_{E\in\mathcal{T}_h}\|\nabla u\|_{L^1(E)}. \tag{3.51}$$

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

$$\|Du\|(B) \leq \sum_{e\in\mathcal{E}_h} |e|^{1/2}\ \|[\![u]\!]\|_{0,e} + \sum_{E\in\mathcal{T}_h} |E|^{1/2}\ \|\nabla u\|_{0,E}. \tag{3.52}$$

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

$$\|Du\|^2(B) \leq 2\left[\sum_{e\in\mathcal{E}_h} |e|^{1/2}\|[\![u]\!]\|_{0,e}\right]^2 + 2\left[\sum_{E\in\mathcal{T}_h} |E|^{1/2}\|\nabla u\|_{0,E}\right]^2. \tag{3.53}$$

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

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

which, by Lemma 3.3, implies

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

with C independent of h. From (2.8),

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

By hypothesis, $u \in V_h + (H^1_0(B))^d$; 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 Theorem 1, [25, p. 183] (last line) we have the equivalence

$$\|D\tilde{u}\|(\mathbb{R}^d) = \|Du\|(B). \tag{3.54}$$

Thus, by applying the Poincaré inequality for BV, Theorem 3.21, we obtain

$$\|u\|_{L^{d/(d-1)}(B)} = \|\tilde{u}\|_{L^{d/(d-1)}(\mathbb{R}^d)} \leq C\ \|D\tilde{u}\|(\mathbb{R}^d) = C\ \|Du\|(B) \leq C\ \|\!|u|\!\| \tag{3.55}$$

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

$$\|v_h\|_{BV} \leq C\, h^{m-1} \tag{3.56}$$

Proof. Apply Theorem 3.20 together with Theorem 3.22. □

4. Final Remarks. Optimal convergence of a stabilized, discontinuous Galerkin 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, elasto-plasticity and fracture.

Acknowledgments. P.N. and D.S. acknowledge the kind hospitality of the Graduate Aeronautical Laboratories during their visits.

REFERENCES

[1] T. H. PIAN AND P. TONG, *Basis of Finite Element Methods for Solid Continua*, Int. J. Num. Meths. Engrng., 1 (1969) pp. 3–28.

[2] G. A. BAKER, *Finite Element Methods for Elliptic Equations Using Nonconforming Elements*, Math. Comp., 31 (1977) pp. 45–59.

[3] J. DOUGLAS, JR., *H^1-Galerkin Method for a Nonlinear Dirichlet Problem*, in Mathematical Aspects of the Finite Element Methods, Springer Lecture Notes in Mathematics, Rome, 606, 1977, pp. 64–86.

[4] J. DOUGLAS, JR. AND T. DUPONT, *Interior Penalty Procedures for Elliptic and Parabolic Galerkin Method*, Lecture Notes in Physics, Springer-Verlag, Berlin, 58, 1976.

[5] J. NITSCHE, *Über ein Variationsprinzip zur Löung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind.*, Abh. Math. Sem. Univ. Hamburg, 36 (1971) pp. 9–15 .

[6] P. PERCELL AND M. F. WHEELER, *A Local Residual Finite Element Procedure for Elliptic Equations*, SIAM J. Numer. Anal., 15 (1978), pp. 705-714.

[7] M. F. WHEELER, *An Elliptic Collocation-Finite Element Method with Interior Penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161 .

[8] D. N. ARNOLD, *An Interior Penalty Finite Element Method with Discontinuous Element*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[9] DOUGLAS N. ARNOLD, FRANCO BREZZI, BERNARDO COCKBURN, AND L. DONATELLA MARINI, *Unified Analysis of Discontinous Galerkin Methods for Elliptic Problems*, SIAM J. Numer. Anal., 39 (2002) pp. 1749-1779.

[10] B. COCKBURN AND C. W. SHU, *The Local Discontinuous Galerkin Method for Time-Dependent Convection-Diffusion Problems*, SIAM J. Numer. Anal., 35 (1998) pp. 2440–2463.

[11] B. COCKBURN AND C. W. SHU, *Runge-Kutta Discontinuous Galerkin Methods*, J. Sci. Comput., 16 (2001) pp. 173–261.

[12] J. TINSLEY ODEN, IVO BABŬSKA, AND CARLOS ERIK BAUMAN, *A Discontinuous HP Finite Element Method for Diffusion Problems*, J. Comput. Phys., 146 (1998) pp. 491–519.

[13] 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) pp. 1676–1706.

[14] PAUL HOUSTON, CHRISTOPH SCHWAB, AND ENDRE SÜLI, *Discontinuous HP-Finite Element Methods for Advection-Diffusion-Reaction Problems*, SIAM J. Numer. Anal., 39 (2002) pp. 2133-2163.

[15] F. BASSI AND S. REBAY, *High-Order Accurate Discontinuous Finite Element Method for the Numerical Solution of the Compressible Navier-Stokes Equations*, J. Comput. Phys., 131 (1997) pp. 267–279.

[16] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo, *Discontinuous Galerkin Approximations for Elliptic Problems*, Numerical Methods for Partial Differential Equations, 16 (2000) pp. 365–378.

[17] F. Brezzi , M. 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, (2001) pp. 197-217.

[18] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and Donatella Marini, *Discontinous Galerkin Methods for Elliptic Problems*, In: Discontinuous Galerkin Methods. Theory, Computation and Applications, B. Cockburn, G.E. Karniadakis, and C.-W. Shu, eds., Lecture Notes in Computational Science and Engineering, 11 (2000) pp. 89-101.

[19] M. Ortiz and A. Pandolfi, *A Class of Cohesive Elements for the Simulation of Three-Dimensional Crack Propagation*, Int. J. Num. Meths. Engrg., 44 (1999) pp.1267-1282.

[20] B. Rivière and M. F. Wheeler, *Optimal Error Estimates for Discontinuous Galerkin Methods Applied to Linear Elasticity Problems*, Texas Institute for Computational and Applied Mathematics Report 00-30, 2000.

[21] P. G. Ciarlet , *The finite element method for elliptic problems*, North-Holland, 1978.

[22] S. C. Brenner and L. Scott Ridgeway, *The Mathematical Theory of Finite Element Methods*, Springer, 1994.

[23] Richard S. Falk, *Nonconforming Finite Element Methods for the Equations of Linear Elasticity*, Math. Comp., 57 (1991) pp. 529-550.

[24] J. E. Roberts and J.-M. Thomas, *Mixed and Hybrid Methods.* In Ciarlet P. .G. and J. L. Lions, Handbook of Numerical Analysis, Vol. II, Elsevier Science Publishers B.V. (North-Holland) 1991.

[25] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, London, 1992.

[26] J. E. Marsden, and T. J. R. Hughes , *Mathematical Foundations of Elasticity.* Prentice Hall, 1983. Reprinted by Dover Publications, NY, 1994.

[27] R. Témam, *Problemes mathématiques en plasticité*, Gauthier-Villars, Paris, 1983.

[28] P. Neff, *On Korn's first inequality with nonconstant coefficients*, Preprint Nr.2080 TU Darmstadt (2000), Proc. Roy. Soc. Edinb., 132A (2002) pp. 221–243.

[29] Philippe G. Ciarlet, *Three-dimensional elasticity*, first ed., Studies in mathematics and its applications, vol. 1, Elsevier, Amsterdam, 1988.

[30] V. Ruas, *Circumventing discrete Korn's inequalities in convergence analyses of nonconforming finite element approximations of vector fields*, ZAMM Z. angew. Math. Mech.76 (1996) pp. 483-484.