Numerische Mathematik manuscript No.
(will be inserted by the editor)

Uniform Convergence and A Posteriori Error Estimators for the Enhanced Strain Finite Element Method

D. Braess[1], C. Carstensen[2], B.D. Reddy[3]

[1] Faculty of Mathematics, Ruhr-University, 44780 Bochum, Germany
[2] Institute for Applied Mathematics and Numerical Analysis, Vienna University of Technology, Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria
[3] Department of Mathematics and Applied Mathematics, University of Cape Town, 7701 Rondebosch, South Africa.

The date of receipt and acceptance will be inserted by the editor

e-mail: `braess@num.ruhr-uni-bochum.de`
e-mail: `Carsten.Carstensen@tuwien.ac.at`
e-mail: `bdr@science.uct.ac.za`

Summary Enhanced strain elements, frequently employed in practice, are known to improve the approximation of standard (non-enhanced) displacement-based elements in finite element computations. The first contribution in this work towards a complete theoretical explanation for this observation is a proof of robust convergence of enhanced element schemes: it is shown that such schemes are locking-free in the incompressible limit, in the sense that the error bound in the a priori estimate is independent of the relevant Lamé constant. The second contribution is a residual-based a posteriori error estimate; the L^2 norm of the stress error is estimated by a reliable and efficient estimator that can be computed from the residuals.

Key words Four-noded quadrilateral, enhanced strains, uniform convergence, a posteriori error estimates

1 Introduction

This work is devoted to rigorous a priori and a posteriori error analyses of enhanced strain finite element schemes for the boundary value problem of linear elasticity. The robust a priori error estimates are derived for a class of triangulations into refined parallelograms and parallelepipeds. Given a piecewise bilinear displacement $\boldsymbol{u}_h$ with affine Green strain $\boldsymbol{\epsilon}(\boldsymbol{u}_h)$, the enhanced strain finite element method is obtained by adding an enhancement $\boldsymbol{a}_h$. The resulting stress approximation reads $\boldsymbol{\sigma}_h = \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h)$ with $\mathcal{C}$ being the elasticity tensor.

Introduced by Simo and Rifai [19], enhanced strain methods are nowadays a popular tool in computational mechanics and are designed to overcome two problems experienced with standard low-order quadrilaterals: (a) they lead to poor results when coarse meshes are used in the solution of bending-dominated problems; and (b) locking is encountered for nearly incompressible materials.

We can prove uniform convergence of the strains, but we cannot prove the same for the spherical part of the stresses. This corresponds to the fact that the constant in the inf-sup condition corresponding to an equivalent mixed method (see (3.8)) is not independent of the Lamé constant λ. The only trial functions available for the stresses are modifications of those derived from the assumed strains. The situation would be different – and more favourable – if, for example, mixed methods with the PEERS elements are used, because of their λ-uniform convergence in practice [8].

Simo and Rifai [19] investigated enhanced strains for plane and geometrically linear problems. The method was subsequently extended to nonlinear problems in two and three dimensions by Simo, Armero, and Taylor [17,18]. Extensive computational studies indicate that the method provides a successful approach for overcoming the aforementioned difficulties (a)-(b).

An analysis of the enhanced strain method has been carried out for affine-equivalent meshes by Reddy and Simo [15]. They established an a priori error estimate which confirms convergence at the standard linear rate,

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \leq c_1 h \|\boldsymbol{f}\|_0. \tag{1.1}$$

Braess [5] has re-examined the sufficient conditions for convergence, in particular relating the stability condition to a strengthened Cauchy inequality, and elucidating the influence of the Lamé constant λ.

The issue of uniform convergence in the incompressible limit, that is, as $\lambda \to \infty$, is not satisfactorily addressed in those works, and is given a proper treatment in this contribution. The a priori analysis of Section 4 guarantees λ-independent asymptotic convergence of the displacement error $\boldsymbol{u} - \boldsymbol{u}_h$ in $V := [H_0^1(\Omega)]^2$ for a class of meshes. The constant c_1 in (1.1) is independent of λ. The arguments involve a stable pair of spaces for an auxiliary Stokes problem, and are thus available for triangulations allowing for an extraction (or filtering) of checkerboard instability modes.

The second contribution in the present paper is the derivation of a residual-based a posteriori error estimate for a wider class of meshes and enhanced elements. The analysis of Section 5 establishes a reliable and efficient error estimate

$$\begin{aligned}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \le c_2 \Bigg(&\sum_{T\in\mathcal{T}} h_T^2 \|\boldsymbol{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_0^2 \\ &+ \sum_{E\in\mathcal{E}} h_E \|[\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_E]\|_{L^2(E)}^2 + \|2\mu \boldsymbol{a}_h\|_0^2 \Bigg)^{1/2} .\end{aligned} \tag{1.2}$$

Here, h_T is the diameter of the element $T \in \mathcal{T}$, and h_E is the length of an edge $E \in \mathcal{E}$; $[\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_E]$ denotes the jump of the normal stress vector across the inner edge E. The constant $c_2 > 0$ depends on the aspect ratio of the elements and on Ω, but neither on the mesh size nor on the Lamé constants λ and μ. The right-hand side of (1.2) may serve as a refinement indicator within an adaptive algorithm. The class of meshes covered by the present a priori error estimates, however, does not include those meshes arising from local mesh refinements.

The rest of the paper is organised as follows. The model problem in linear elasticity with pure homogeneous boundary conditions is introduced in Section 2. The finite element method with enhanced assumed strains is described in Section 3, and a first a priori estimates of the error are presented. Section 4 establishes the λ-independent estimate (1.1). The main argument is the construction of an interpolation operator which preserves the divergence-free property of the displacement. Our a priori estimates are restricted to special grids, while the reliable and efficient a posteriori error estimate (1.2) is addressed in Section 5 for more general meshes and enhanced schemes.

2 Boundary value problem of elasticity

Consider an isotropic linear elastic material body which occupies a bounded domain Ω in $\mathbb{R}^d$ ($d = 2$ or 3) with Lipschitz boundary Γ.

For a prescribed body force $\boldsymbol{f}$, the governing equilibrium equation in Ω reads

$$-\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{f} \tag{2.1}$$

for the symmetric Cauchy stress tensor $\boldsymbol{\sigma}$. The infinitesimal strain tensor $\boldsymbol{\epsilon}$ is defined by

$$\boldsymbol{\epsilon}(\boldsymbol{u}) := \tfrac{1}{2}(\nabla \boldsymbol{u} + [\nabla \boldsymbol{u}]^T) \tag{2.2}$$

for the displacement vector $\boldsymbol{u}$. It is assumed to satisfy the homogeneous Dirichlet boundary condition

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on} \quad \Gamma.$$

With the fourth-order elasticity tensor $\mathcal{C}$, the elastic constitutive equation reads

$$\boldsymbol{\sigma} = \mathcal{C}\boldsymbol{\epsilon} = \lambda(\operatorname{tr} \boldsymbol{\epsilon})\mathbf{1} + 2\mu\, \boldsymbol{\epsilon}, \tag{2.3}$$

and has the inverse relation

$$\boldsymbol{\epsilon} = \mathcal{C}^{-1}\boldsymbol{\sigma} = \frac{1}{2\mu}\boldsymbol{\sigma} - \frac{\lambda}{2\mu(d\lambda + 2\mu)}(\operatorname{tr} \boldsymbol{\sigma})\mathbf{1}. \tag{2.4}$$

Here, $\mathbf{1}$ is the identity, and λ and μ are the homogeneous Lamé constants. Pointwise stability of $\mathcal{C}$ is clear for positive Lamé constants λ and μ [14]. This work is devoted to the incompressible limit $\lambda \to \infty$.

We will make use of the space $L^2(\Omega)$ of square-integrable functions defined on Ω with the inner product and norm being denoted by $(\cdot,\cdot)_0$ and $\|\cdot\|_0$, respectively. We recall also the definition of the Sobolev spaces $H^m(\Omega)$, with m being a non-negative integer, as equivalence classes of functions with generalised derivatives of order $\leq m$ in $L^2(\Omega)$. The Sobolev spaces are Hilbert spaces with inner product and associated norm

$$(u,v)_m := \int_\Omega \sum_{|\boldsymbol{\alpha}|\leq m} D^{\boldsymbol{\alpha}} u(\boldsymbol{x}) D^{\boldsymbol{\alpha}} v(\boldsymbol{x}) \; dx \quad \text{and} \quad \|v\|_m := (v,v)_m^{1/2}. \tag{2.5}$$

Here $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ is a multi-index whose components α_j are nonnegative integers, $|\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_d$, and as usual $D^{\boldsymbol{\alpha}} = \partial^{|\boldsymbol{\alpha}|}/\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$. The seminorm $|\cdot|_m$ on $H^m(\Omega)$ is defined by

$$|v|_m^2 := \int_\Omega \sum_{|\boldsymbol{\alpha}|=m} D^{\boldsymbol{\alpha}} v(\boldsymbol{x}) D^{\boldsymbol{\alpha}} v(\boldsymbol{x}) \; dx. \tag{2.6}$$

The space $H_0^1(\Omega)$ consists of functions in $H^1(\Omega)$ which vanish on the boundary in the sense of traces. The space $H^{-1}(\Omega)$ is the dual space of $H_0^1(\Omega)$. Finally, we denote by $L_0^2(\Omega)$ the subspace of functions in L^2 with zero mean; that is,

$$L_0^2(\Omega) := \left\{ v \in L^2(\Omega) : \int_\Omega v\,dx = 0 \right\}. \tag{2.7}$$

We now define the standard variational problem in linear elasticity. For this purpose, $V := [H_0^1(\Omega)]^d$ will be the space of admissible displacements, V' its dual, and we define the bilinear form $a(\cdot,\cdot)$ and linear functional $\ell(\cdot)$ by

$$a : V \times V \to \mathbb{R}, \qquad a(\boldsymbol{u},\boldsymbol{v}) := \int_\Omega \mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{u}) : \boldsymbol{\epsilon}(\boldsymbol{v})\ dx, \tag{2.8}$$

$$\ell : V \to \mathbb{R}, \qquad \ell(\boldsymbol{v}) := \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\ dx. \tag{2.9}$$

Problem S. *Given $\ell \in V'$, find $\boldsymbol{u} \in V$ that satisfies*

$$a(\boldsymbol{u},\boldsymbol{v}) = \ell(\boldsymbol{v}) \quad \textit{for all} \quad \boldsymbol{v} \in V. \tag{2.10}$$

The present assumptions on $\mathcal{C}$ guarantee that $a(\cdot,\cdot)$ is symmetric, continuous, and V-elliptic: there exist positive constants c_3 and $c_4(\lambda)$, the latter being dependent on λ, such that, for all $\boldsymbol{u}$, $\boldsymbol{v} \in V$,

$$|a(\boldsymbol{u},\boldsymbol{v})| \le c_3 \|\boldsymbol{u}\|_1 \|\boldsymbol{v}\|_1, \quad \text{and} \quad c_4(\lambda)\|\boldsymbol{v}\|_1^2 \le a(\boldsymbol{v},\boldsymbol{v}). \tag{2.11}$$

It is known (see, for example, [4]) that Problem S has a unique solution $\boldsymbol{u}$ and

$$\|\boldsymbol{u}\|_1 \le 1/c_4(\lambda)\, \|\boldsymbol{f}\|_{-1}\,. \tag{2.12}$$

3 Enhanced strain finite element approximations

In this section we introduce the finite element space with enhanced assumed strains on a polygonal or polyhedral domain for $d = 2$ or $d = 3$, respectively. Let $\mathcal{T}_h$ be a regular triangulation of parallelograms (resp. parallelipipeds) on Ω with mesh parameter $h := \max_{T \in \mathcal{T}_h} h_T$. A typical element T in $\mathcal{T}_h$ is generated by an affine map F from the reference element $(-1,1)^d$. Let $\mathcal{Q}_1(T)$ be the space of bilinear (resp. trilinear) polynomials on T; that is, polynomials p in $\mathcal{Q}_1$ have the form $p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}| \le 1} a_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$. Set

$$V^h := \{\boldsymbol{v}_h \in C(\overline{\Omega})^d : \ \boldsymbol{v}_h = \boldsymbol{0} \text{ on } \Gamma,\ \boldsymbol{v}_h|_T \in \mathcal{Q}_1(T) \text{ for all } T \in \mathcal{T}_h\}. \tag{3.1}$$

We define the discrete strain $\boldsymbol{\epsilon}_h$ by

$$\boldsymbol{\epsilon}_h = \boldsymbol{\epsilon}(\boldsymbol{v}_h) + \boldsymbol{b}_h; \tag{3.2}$$

that is, the strain comprises the conventional or consistent part $\boldsymbol{\epsilon}(\boldsymbol{v}_h)$ by (2.2), and, in addition, the enhanced assumed strain $\boldsymbol{b}_h$. The finite element spaces $\tilde{E}^h$ for enhanced strains are subspaces of the space Λ of symmetric $d \times d$ matrix-valued functions in $L_{2,0}(\Omega)$: that is,

$$\begin{aligned} \tilde{E}^h \subset \Lambda &:= L_{2,0}(\Omega; \mathbb{R}^{d\times d}_{sym}) \\ &:= \{\boldsymbol{b} \in L^2(\Omega; \mathbb{R}^{d\times d}_{sym}) : \; C_{ijkl} b_{kl} \in L_{2,0}(\Omega)\}. \end{aligned}$$

Problem $\mathbf{E}_h$. *Find $(\boldsymbol{u}_h, \boldsymbol{a}_h) \in V^h \times \tilde{E}^h$ such that, for all $(\boldsymbol{v}_h, \boldsymbol{b}_h) \in V^h \times \tilde{E}^h$,*

$$\int_\Omega \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h) + \boldsymbol{b}_h) \; dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h \; dx. \tag{3.3}$$

Throughout the a priori error analysis in Section 4 we focus on the following standard example, but comment on other possibilities in Remark 1. In two dimensions, the original choice of Simo and Rifai [19] reads

$$\begin{aligned} \tilde{E}^h &:= \{\boldsymbol{b}_h \in \Lambda : \boldsymbol{b}_h|_T \in \mathcal{L}(T) \quad \text{for all } T \in \mathcal{T}_h\}, \\ \mathcal{L}(T) &:= \operatorname{span} \left\{ \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & y \end{pmatrix}, \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}, \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} \right\}. \end{aligned} \tag{3.4}$$

Here, x and y are the (local) coordinates in the reference element.

Remark 1. The method introduced by Simo and Rifai is advantageous also in the case of *shear locking*. In this paper we restrict our attention to effects which are related to *volume locking*; in this context, the set $\tilde{E}^h$ may be reduced to the first two contributions of $\mathcal{L}(T)$ in (3.4). That is, $\tilde{E}^h$ may be replaced by

$$\{\boldsymbol{b}_h \in \Lambda : \boldsymbol{b}_h|_T \in \operatorname{span}\{\operatorname{diag}(x,0), \operatorname{diag}(0,y)\} \quad \text{for all } T \in \mathcal{T}_h\}. \tag{3.5}$$

This reduction does not require any changes in the analysis that follows.

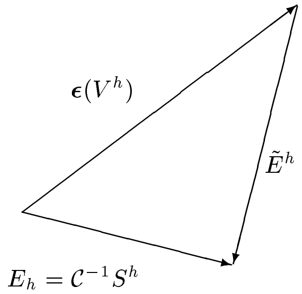


Figure 1. $\mathcal{C}$-orthogonal decomposition of $\boldsymbol{\epsilon}(V^h)$ to illustrate the relationship between the spaces V^h, $\tilde{E}^h$, and S^h through $\mathcal{C}^{-1}\boldsymbol{\sigma}_h = \boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h$.

There is another reduction in the same spirit. It suffices that the enhanced strains apply only to the volumetric part of the stored energy. Thus, the variational problem (3.3)–(3.4) may be replaced by

$$2\mu \int_\Omega \boldsymbol{\epsilon}(\boldsymbol{u}_h) : \boldsymbol{\epsilon}(\boldsymbol{v}_h)\ dx + \lambda \int_\Omega (\operatorname{div} \boldsymbol{u}_h + \operatorname{tr} \boldsymbol{a}_h)\,(\operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h)\, dx$$
$$= \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h\ dx \quad \text{for all } (\boldsymbol{v}_h, \boldsymbol{b}_h) \in V^h \times \tilde{E}^h. \tag{3.6}$$

The reductions above are of interest when nonlinear problems are treated. Then the softening induced by the enhanced strains is to be kept as small as possible since it is reported [16,19] that numerical solutions may exhibit instabilities such as checkerboard modes in the incompressible limit. Whether such effects are possible for linear problems, in unclear from the existing literature. □

Problem E_h is equivalent to a mixed method [5,23], [4, p.152] and was in fact first posed as a mixed problem [19]. The corresponding discrete stress space S^h consists of all $\boldsymbol{\tau}_h$ in $\mathcal{C}(\boldsymbol{\epsilon}(V^h) + \tilde{E}^h)$ that are L^2 orthogonal to $\tilde{E}^h$, i.e.,

$$S^h \subset \mathcal{C}(\boldsymbol{\epsilon}(V^h) + \tilde{E}^h) \quad \text{with} \quad S^h \perp \tilde{E}^h. \tag{3.7}$$

Theorem 1 *[5, 23] Problem E_h is equivalent to the following saddle point problem:* Find $(\boldsymbol{u}_h, \boldsymbol{\sigma}_h) \in V^h \times S^h$ such that

$$\begin{aligned} (\mathcal{C}^{-1}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h)_0 - (\boldsymbol{\tau}_h, \boldsymbol{\epsilon}(\boldsymbol{u}_h))_0 &= 0 && \text{for all } \boldsymbol{\tau}_h \in S^h\,, \\ (\boldsymbol{\sigma}_h, \boldsymbol{\epsilon}(\boldsymbol{v}_h))_0 &= (\boldsymbol{f}, \boldsymbol{v}_h)_0 && \text{for all } \boldsymbol{v}_h \in V^h. \end{aligned} \tag{3.8}$$

Here $(\cdot,\cdot)_0$ denotes the L^2 inner product, for scalar-, vector-, or matrix-valued functions on Ω.

Proof. Let $\boldsymbol{u}_h$, $\boldsymbol{a}_h$ be a solution of (3.3) and set $\boldsymbol{\sigma}_h := \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h)$. It follows from (3.3) that $(\boldsymbol{\sigma}_h, \boldsymbol{b}_h)_0 = 0$ for all $\boldsymbol{b}_h \in \tilde{E}^h$. We conclude from (3.7) that $\boldsymbol{\sigma}_h \in S^h$. Moreover, (3.3) implies $(3.8)_2$, and $(3.8)_1$ is a consequence of the orthogonality relation in (3.7). □

The variational formulation (3.8) is also obtained when a discretization is based on the Hellinger-Reissner principle.

Remark 2. The space S^h associated with (3.4) is easily described for rectangular grids [2]. It can be chosen as those stresses which have on each rectangle the form

$$\tau_{11} = \alpha_{11} + \gamma_{11}\, y, \quad \tau_{12} = \alpha_{12}, \quad \text{and } \tau_{22} = \alpha_{22} + \beta_{22}\, x. \tag{3.9}$$

Remark 3. There is also another interpretation of the mixed formulation (3.8). The displacement vectors $\boldsymbol{v}_h \in V^h$ lead to strains with piecewise linear traces. The enhancements specified above ensure that the L^2 projections onto piecewise constant traces are included. This causes a softening of the energy function that is appropriate for $\lambda \to \infty$. The internal energy is not directly taken from the strain but from its projection onto $\mathcal{C}^{-1}S^h$. The projection is orthogonal with respect to the energy norm. Although this property assists in understanding the softening influence of the enhanced strains (cf. [4, Chapter III.5]), the dependence of $\mathcal{C}^{-1}S^h$ on the Lamé parameter λ is a complication of its applicability.

The pairing S^h and V^h must satisfy an inf-sup condition in order to guarantee stability of the formulation (3.8). This corresponds to a strengthened Cauchy inequality for the finite element spaces $\boldsymbol{\epsilon}(V^h)$ and $\tilde{E}^h$ with respect to the energy norm [4, Chapter III.5].

Condition (a) in the following lemma is called a *strengthened Cauchy inequality.*

Lemma 1 ([5, Lemma A]) *Let X be a Hilbert space with inner product $(\cdot,\cdot)$ with closed subspaces V and W. Let the constant c_5 satisfy $0 < c_5 < 1$. Then, the following assertions (a), (b), and (c) are pairwise equivalent:*

(a) $(v, w) \le c_5 \|v\|\,\|w\|$ *for all* $v \in V$ *and* $w \in W$*;*

(b) $\|v + w\| \ge \sqrt{1 - c_5^2}\, \|v\|$ *for all* $v \in V$ *and* $w \in W$*;*

(c) $\|v + w\| \ge \sqrt{\tfrac{1}{2}(1 - c_5)}\, (\|v\| + \|w\|)$ *for all* $v \in V$ *and* $w \in W$*.* □

The spaces V^h and S^h defined by (3.1) and (3.4) satisfy a strengthened Cauchy inequality. Specifically,

$$(\nabla \boldsymbol{v}_h, \boldsymbol{b}_h)_0 \leq c_5 \|\nabla \boldsymbol{v}_h\|_0 \, \|\boldsymbol{b}_h\|_0 \quad \text{for all } \boldsymbol{v}_h \in V^h \text{ and } \boldsymbol{\eta}_h \in \tilde{E}^h \tag{3.10}$$

was established with $c_5 < 1$ in [5,15]. The Lamé constant λ is not present in the constant c_5 since the finite element spaces and the L^2 norm are independent of λ by definition.

Remark 4. A uniform strengthened Cauchy inequality does not hold for the energy norm. That is, the inequality

$$\begin{aligned} (\mathcal{C}\nabla \boldsymbol{v}_h, \boldsymbol{b}_h)_0 \leq c_6(\lambda) \|\mathcal{C}^{1/2}\nabla \boldsymbol{v}_h\|_0 \, \|\mathcal{C}^{1/2}\boldsymbol{b}_h\|_0 \\ \text{for all } \boldsymbol{v}_h \in V^h \text{ and } \boldsymbol{b}_h \in \tilde{E}^h \end{aligned} \tag{3.11}$$

is not valid for c_6 independent of λ. Here, V^h, $\tilde{E}^h \subset \Lambda$ and Λ is endowed with the energy scalar product $(\mathcal{C}\cdot, \cdot)_0$. Lemma 1 permits a heuristic interpretation for a degenerating constant $c_6(\lambda)$ in case of the energy norm: The locking of the $\mathcal{Q}_1$ element has its origin in the fact that

$$\lambda^{1/2} \, \| \operatorname{div} \boldsymbol{v}_h \|_0$$

is too large for most functions [5, Theorem 3]. On the other hand, by adding an appropriate enhanced strain $\boldsymbol{b}_h$,

$$\lambda^{1/2} \, \| \operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h \|_0$$

becomes small. It follows from Lemma 1(b) that this is only possible if the constant $c_6(\lambda)$ is close to 1. Hence, $c_6(\lambda) \to 1$ holds for $\lambda \to \infty$.

The failure of the uniform strengthened Cauchy inequality (3.11) is proved as follows: Let p^{ch} be the scalar checkerboard mode function (generated by the functions in Fig. 2d). Set

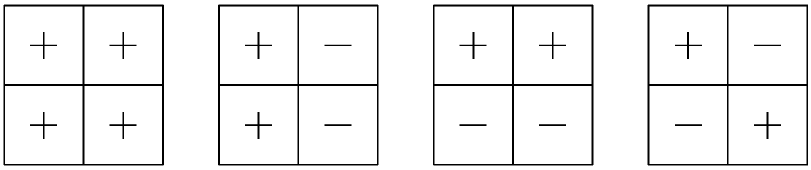
$$v_{h,x}|_T := p^{ch} \, xy \quad \text{in each quadrilateral } T$$

where x and y are local coordinates as in (3.4) and $v_{h,y} := 0$. Then we have

$$\operatorname{div} \boldsymbol{v}_h|_T = p^{ch} \, y \quad \text{in each quadrilateral } T,$$

and $\operatorname{div} \boldsymbol{v}_h \in \operatorname{tr} \tilde{E}^h$. If $\lambda \to \infty$, the volume term dominates the energy functional, and a nontrivial constant $c_5(\lambda)$ in the strengthened Cauchy inequality (3.11) cannot be independent of λ and h.

We will make use of the strengthened Cauchy inequality for the L^2 norm as in [5], or in a slightly different context in [13].



(a) (b) (c) (d)

Figure 2 a-d. Restrictions of the basis functions of M_h to the macro-elements with values ± 1 according to the sign indicated inside the elements.

Lemma 2 *Assume that the spaces $\boldsymbol{\epsilon}(V^h)$ and $\tilde{E}^h$ satisfy a strengthened Cauchy inequality (3.10). Then, Problem E_h has a unique solution $(\boldsymbol{u}_h, \boldsymbol{a}_h) \in V^h \times \tilde{E}^h$ and*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_1 + \|\boldsymbol{a}_h\|_0 \leq c_7 \left(\inf_{\boldsymbol{v}_h \in V^h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_1 + c_8 h \, \|\boldsymbol{\epsilon}(\boldsymbol{u})\|_0 \right). \quad (3.12)$$

The constant c_7 is (λ, μ)-independent and $c_8 = (1 + \lambda/\mu) \frac{2}{c_4(1-c_5)}$.
□

Although not stated in this form, Lemma 2 was first proved by Reddy and Simo [15]; the explicit form of the constant c_8 is due to Braess [5]. It is clear from the expression for c_8 that (3.12) does not guarantee uniform convergence in the incompressible limit $\lambda \to \infty$. The remedy in the next section will be more involved.

4 Uniform convergence in the incompressible limit

This section is devoted to an a priori estimate which is uniform in λ as $\lambda \to \infty$. For convenience, we confine our attention to plane problems. Let Ω be a polygonal domain and suppose that $\mathcal{T}_{2h}$ is a regular triangulation of Ω into parallelograms. Further let $\mathcal{T}_h$ be a refinement of $\mathcal{T}_{2h}$, i.e., each parallelogram $T \in \mathcal{T}_{2h}$ is partitioned into four congruent sub-parallelograms to generate $\mathcal{T}_h$.

Theorem 2 *There exists a constant $c_9 > 0$ independent of λ, $\mathcal{T}_h$, $\boldsymbol{f}$, and $\boldsymbol{u}$ such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_1 + \|\boldsymbol{a}_h\|_0 \leq c_9 h \|\boldsymbol{f}\|_0. \quad (4.1)$$

The essential point in the proof will be the design of an interpolant to the solution of the continuous problem for which the trace of the discrete strain,

$$\operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h,$$

is small. We will deduce this from the fact that a $\mathcal{Q}_1$ finite element is constructed for which the mean value of $\operatorname{div} \boldsymbol{v}_h$ in each element is small. Set

$$\operatorname{div_h} \boldsymbol{w}_h := \operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h \quad \text{for each } \boldsymbol{w}_h = (\boldsymbol{v}_h, \boldsymbol{b}_h) \in V^h \times \tilde{E}^h. \tag{4.2}$$

The crucial tool is a lemma hidden behind the analysis in [3, p. 231] and based on a regularity result in [22] plus a property of the divergence operator [1].

Lemma 3 *Let Ω be a polygonal domain. Assume that there is a linear operator*

$$\Pi_h : (H^2(\Omega) \cap H_0^1(\Omega))^2 \longrightarrow V^h \times \tilde{E}^h$$

such that, for all $\boldsymbol{v} \in (H^2(\Omega) \cap H_0^1(\Omega))^2$, we have

$$\|(\boldsymbol{v}, 0) - \Pi_h \boldsymbol{v}\|_{H^1(\Omega)^2 \times L^2(\Omega)^{2\times 2}} \leq c_{10}\, h \|\boldsymbol{v}\|_2 \tag{4.3}$$

and

$$\operatorname{div_h}(\Pi_h \boldsymbol{v}) = 0 \quad \textit{whenever} \quad \operatorname{div} \boldsymbol{v} = 0. \tag{4.4}$$

If $\boldsymbol{u} \in V$ is the solution of the variational problem (2.1)–(2.3), we have

$$\lambda \| \operatorname{div} \boldsymbol{u} - \operatorname{div_h} \Pi_h \boldsymbol{u}\|_0 \leq c_{11}\, h \|\boldsymbol{f}\|_0. \tag{4.5}$$

Proof For homogeneous boundary conditions, a regularity result due to Vogelius [22] shows that

$$\|\boldsymbol{u}\|_2 + \lambda \| \operatorname{div} \boldsymbol{u}\|_1 \leq c_{12} \|\boldsymbol{f}\|_0. \tag{4.6}$$

Since Ω is polygonal, Theorem 3.1 in [1] asserts the existence of $\boldsymbol{u}_1 \in (H^2(\Omega) \cap H_0^1(\Omega))^2$ with

$$\operatorname{div} \boldsymbol{u}_1 = \operatorname{div} \boldsymbol{u} \quad \text{and} \quad \|\boldsymbol{u}_1\|_2 \leq c_{13} \| \operatorname{div} \boldsymbol{u}\|_1. \tag{4.7}$$

By (4.3), (4.6), and (4.7),

$$\begin{aligned} \|\operatorname{div_h}((\boldsymbol{u}_1, 0) - \Pi_h \boldsymbol{u}_1)\|_0 &\leq c_{10} h \|\boldsymbol{u}_1\|_2 \leq c_{10} c_{13} h \| \operatorname{div} \boldsymbol{u}\|_1 \\ &\leq \frac{c_{10} c_{13} h}{\lambda} \|\boldsymbol{f}\|_0. \end{aligned} \tag{4.8}$$

From (4.4) and $\operatorname{div}(\boldsymbol{u} - \boldsymbol{u}_1) = 0$ it follows that $\operatorname{div_h}(\Pi_h(\boldsymbol{u} - \boldsymbol{u}_1)) = 0$ and

$$\operatorname{div_h}((\boldsymbol{u}, 0) - \Pi_h \boldsymbol{u}) = \operatorname{div_h}((\boldsymbol{u}_1, 0) - \Pi_h \boldsymbol{u}_1).$$

Hence, (4.8) reads

$$\|\operatorname{div_h}((\boldsymbol{u}, 0) - \Pi_h \boldsymbol{u})\|_0 \leq \frac{c_{10} c_{13} h}{\lambda} \|\boldsymbol{f}\|_0. \quad \square \tag{4.9}$$

Construction 1. We establish an interpolation operator Π_h with the properties (4.3)–(4.4) in two steps, by considering an appropriate Stokes problem.

Step 1. Let $M_h \subset L_0^2(\Omega)$ be the set of piecewise constant functions with respect to the grid $\mathcal{T}_h$. The aim is to construct a function $\boldsymbol{v}_h \in V^h$ such that

$$(\operatorname{div} \boldsymbol{v}_h, q_h)_0 = (\operatorname{div} \boldsymbol{v}, q_h)_0 \quad \text{for all} \quad q_h \in M_h\,, \tag{4.10}$$

at least if $\operatorname{div} \boldsymbol{v} = 0$. Since the Q_1-P_0 finite element is known to suffer from checkerboard instabilities, we apply a filter as in [11, p.167]. The restrictions of the functions in M_h to a macro-element are spanned by the four functions depicted in Fig. 2 a-d. Let A_h be the subspace of M_h spanned by the checkerboard modes shown in Fig. 2d (on all macro elements), and set

$$\begin{aligned} \tilde{V^h} &:= \{v_h \in V^h : \ (\operatorname{div} \boldsymbol{v}_h, q_h) = 0 \quad \text{for all } q_h \in A_h\}, \\ \tilde{M}_h &:= M_h \cap A_h^\perp. \end{aligned} \tag{4.11}$$

The pair $(\tilde{V}_h, \tilde{M}_h)$ is known to be stable for the Stokes problem (see, for example, [11, p.167]).

Given $\boldsymbol{v} \in (H^2(\Omega) \cap H_0^1(\Omega))^2$, we perform Fortin interpolation; that is, we choose $(\boldsymbol{v}_h, p_h) \in \tilde{V^h} \times \tilde{M}_h$ by

$$\begin{aligned} (\nabla \boldsymbol{v}_h, \nabla \boldsymbol{z}_h)_0 + (\operatorname{div} \boldsymbol{z}_h, p_h)_0 &= (\nabla \boldsymbol{v}, \nabla \boldsymbol{z}_h)_0 \qquad \text{for all} \quad \boldsymbol{z}_h \in \tilde{V^h}, \\ (\operatorname{div} \boldsymbol{v}_h, q_h)_0 \qquad\qquad &= (\operatorname{div} \boldsymbol{v}, q_h)_0 \qquad \text{for all } q_h \in \tilde{M}_h. \end{aligned} \tag{4.12}$$

The resulting function $\boldsymbol{v}_h$ satisfies (4.10) whenever $\operatorname{div} \boldsymbol{v} = 0$. Indeed, the relations that are not covered by $(4.12)_2$ are directly built into the formulation of the space $\tilde{V^h}$ by $(4.11)_1$.

An error estimate is obtained in the usual way. Let $\boldsymbol{v}_{2h,0} \in V^{2h} \subset \tilde{V^h}$ be the nodal interpolant to $\boldsymbol{v}$. Standard scaling arguments combined with the Bramble–Hilbert lemma lead to $\|\boldsymbol{v} - \boldsymbol{v}_{2h,0}\|_1 \le c_{14} h |\boldsymbol{v}|_2$. Hence,

$$\|\operatorname{div}(\boldsymbol{v} - \boldsymbol{v}_{2h,0})\|_0 \le \sqrt{2}\,\|\boldsymbol{v} - \boldsymbol{v}_{2h,0}\|_1 \le c_{14}\sqrt{2}\, h |\boldsymbol{v}|_2. \tag{4.13}$$

By subtracting $(\nabla \boldsymbol{v}_{2h,0}, \nabla \boldsymbol{z}_h)_0$ and $(\operatorname{div} \boldsymbol{v}_{2h,0}, q_h)_0$, resp., on both sides of (4.12) we obtain a discrete Stokes problem for $\boldsymbol{v}_h - \boldsymbol{v}_{2h,0}$. The stability of the Stokes problem asserts that

$$\|\boldsymbol{v}_h - \boldsymbol{v}_{2h,0}\|_1 + \|\operatorname{div}(\boldsymbol{v}_h - \boldsymbol{v}_{2h,0})\|_0 \le c_{15} \|\boldsymbol{v} - \boldsymbol{v}_{2h,0}\|_1 \le c_{14} c_{15} h\, |\boldsymbol{v}|_2$$

and with the triangle inequality

$$\|\boldsymbol{v}_h - \boldsymbol{v}\|_1 + \|\operatorname{div}(\boldsymbol{v}_h - \boldsymbol{v})\|_0 \leq c_{14}(\sqrt{2} + c_{15})h\,|\boldsymbol{v}|_2. \tag{4.14}$$

Step 2. Let $\pi_h : L^2(\Omega) \to M_h$ be the L^2 orthogonal projection. Having $\boldsymbol{v}_h$ from the construction in Step 1, we will construct $\boldsymbol{b}_h \in \tilde{E}^h$ such that

$$\begin{aligned} \operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h &= \pi_h(\operatorname{div} \boldsymbol{v}_h), \\ \epsilon_{12}(\boldsymbol{v}_h) + b_{h,12} &= \pi_h(\epsilon_{12}(\boldsymbol{v}_h)). \end{aligned} \tag{4.15}$$

As $\boldsymbol{v}_h \in V^h$, all components of $\boldsymbol{\epsilon}(\boldsymbol{v}_h)$ are affine functions on each $T \in \mathcal{T}_h$, and we find $\boldsymbol{b}_h \in \mathcal{L}(T)$ from (3.4) such that

$$\operatorname{div} \boldsymbol{v}_h + \operatorname{tr} \boldsymbol{b}_h \text{ and } \epsilon_{12} + b_{h,12} \text{ are constant on each } T \in \mathcal{T}_h, \tag{4.16}$$

that is, constant on each quadrilateral in the fine grid. Moreover, the mean value of $\boldsymbol{b}_h$ on each $T \in \mathcal{T}_h$ vanishes. This observation and (4.16) indeed imply (4.15).

Now we set

$$\Pi_h(\boldsymbol{v}) := (\boldsymbol{v}_h, \boldsymbol{b}_h) \in V^h \times \tilde{E}^h. \tag{4.17}$$

Notice that, in the case $\operatorname{div} \boldsymbol{v} = 0$, (4.10) implies that $\operatorname{div} \boldsymbol{v}_h \perp M_h$, and the mean value of $\operatorname{div} \boldsymbol{v}_h$ vanishes on each element. Hence, $\operatorname{div}_h \Pi_h(\boldsymbol{v}) = \pi_h(\operatorname{div} \boldsymbol{v}_h) = 0$, and (4.4) holds.

From the usual approximation argument combined with (4.14)–(4.15) we conclude that

$$\begin{aligned} \|\operatorname{tr} \boldsymbol{b}_h\|_0 &= \|\pi_h(\operatorname{div} \boldsymbol{v}_h) - \operatorname{div} \boldsymbol{v}_h\|_0 \\ &= \|\pi_h(\operatorname{div} \boldsymbol{v}_h - \operatorname{div} \boldsymbol{v}) + (\pi_h(\operatorname{div} \boldsymbol{v}) - \operatorname{div} \boldsymbol{v}) + \operatorname{div}(\boldsymbol{v} - \boldsymbol{v}_h)\|_0 \\ &\leq 2\|\operatorname{div} \boldsymbol{v}_h - \operatorname{div} \boldsymbol{v}\|_0 + c_{16}h\|\operatorname{div} \boldsymbol{v}\|_1 \leq c_{17}h\|\boldsymbol{v}\|_2 \,. \end{aligned}$$

Similarly, the off-diagonal entries of $\boldsymbol{b}_h$ can be estimated, and it follows that

$$\|\boldsymbol{b}_h\|_0 \leq 4c_{17}h\|\boldsymbol{v}\|_2 \,. \qquad \square \tag{4.18}$$

Remark 5. When $\tilde{E}^h$ is the smaller set suggested in (3.5), then $(4.10)_2$ can be ignored during the construction. This will indeed be appropriate for verifying Theorem 2 for the modified set (3.5).

Proof of Theorem 2. We start the proof in the spirit of the lemma of Berger, Scott, and Strang that is often quoted as the second Strang lemma, but later we will make use of special properties.

For convenience, we introduce the bilinear form

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta})_{\mathcal{C}} := \int_\Omega \boldsymbol{\epsilon} : \mathcal{C}\boldsymbol{\eta}\, dx. \tag{4.19}$$

The original Problem S may be formulated as

$$(\boldsymbol{\epsilon}(\boldsymbol{u}),\boldsymbol{\epsilon}(\boldsymbol{v}))_{\mathcal{C}} = (\boldsymbol{f},\boldsymbol{v})_0 \qquad \text{for all } \boldsymbol{v} \in H_0^1(\Omega), \tag{4.20}$$

while the discretisation with Problem E_h may be rewritten as

$$\begin{aligned} (\boldsymbol{\epsilon}(\boldsymbol{u}_h)+\boldsymbol{a}_h,\boldsymbol{\epsilon}(\boldsymbol{v}))_{\mathcal{C}} &= (\boldsymbol{f},\boldsymbol{v})_0 && \text{for all } \boldsymbol{v}\in V^h,\\ (\boldsymbol{\epsilon}(\boldsymbol{u}_h)+\boldsymbol{a}_h,\boldsymbol{\eta})_{\mathcal{C}} &= 0 && \text{for all } \boldsymbol{\eta}\in \tilde{E}^h. \end{aligned} \tag{4.21}$$

Now we make use of Construction 1. Let

$$(\boldsymbol{w}_h,\boldsymbol{b}_h) := \Pi_h \boldsymbol{u}, \tag{4.22}$$

and set $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{u})$ and $\boldsymbol{\tau}_h := \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{w}_h)+\boldsymbol{b}_h)$. From Lemma 3 and (4.18) we know that

$$\|\boldsymbol{u}-\boldsymbol{w}_h\|_1 + \|\boldsymbol{b}_h\|_0 + \lambda\|\operatorname{div}\boldsymbol{u} - \operatorname{div}\boldsymbol{w}_h - \operatorname{tr}\boldsymbol{b}_h\|_0 \le c_{18}h\|\boldsymbol{f}\|_0 \tag{4.23}$$

and thus

$$\|\boldsymbol{\sigma}-\boldsymbol{\tau}_h\|_0 \le c_{19}h\|\boldsymbol{f}\|_0\,. \tag{4.24}$$

By subtracting $\boldsymbol{\epsilon}(\boldsymbol{w}_h)+\boldsymbol{b}_h$ from the first argument in (4.21) and recalling (4.20), it follows that, for all $\boldsymbol{v}_h\in V^h$ and all $\boldsymbol{\eta}_h\in\tilde{E}^h$,

$$\begin{aligned} (\boldsymbol{\epsilon}(\boldsymbol{u}_h-\boldsymbol{w}_h)+\boldsymbol{a}_h-\boldsymbol{b}_h,\boldsymbol{\epsilon}(\boldsymbol{v}_h))_{\mathcal{C}} &= (\boldsymbol{\epsilon}(\boldsymbol{u}-\boldsymbol{w}_h)-\boldsymbol{b}_h,\boldsymbol{\epsilon}(\boldsymbol{v}_h))_{\mathcal{C}},\\ (\boldsymbol{\epsilon}(\boldsymbol{u}_h-\boldsymbol{w}_h)+\boldsymbol{a}_h-\boldsymbol{b}_h,\boldsymbol{\eta}_h)_{\mathcal{C}} &= -(\boldsymbol{\epsilon}(\boldsymbol{w}_h)+\boldsymbol{b}_h,\boldsymbol{\eta}_h)_{\mathcal{C}}. \end{aligned} \tag{4.25}$$

We look at the right-hand side of $(4.25)_2$ in more detail:

$$\begin{aligned} (\boldsymbol{\epsilon}(\boldsymbol{w}_h)+\boldsymbol{b}_h,\boldsymbol{\eta}_h)_{\mathcal{C}} = 2\mu &\int_\Omega (\epsilon_{11}(\boldsymbol{w}_h)\eta_{h,11}+\epsilon_{22}(\boldsymbol{w}_h)\eta_{h,22})\,dx\\ &+2\mu\int_\Omega (b_{h,11}\eta_{h,11}+b_{h,22}\eta_{h,22})\,dx\\ &+4\mu\int_\Omega (\epsilon_{12}(\boldsymbol{w}_h)+b_{h,12})\,\eta_{h,12}dx\\ &+2\lambda\int_\Omega (\operatorname{div}\boldsymbol{w}_h+\operatorname{tr}\boldsymbol{b}_h)\operatorname{tr}\boldsymbol{\eta}_h\,dx\,. \end{aligned}$$

The first integral vanishes since $\mathcal{L}(\mathcal{T}_h)$ has been chosen such that the functions in the integral are orthogonal on each element. It follows from (4.16) that the third and the fourth integral also vanish. Hence, only the second integral may be nonzero and

$$|(\boldsymbol{\epsilon}(\boldsymbol{w}_h)+\boldsymbol{b}_h,\boldsymbol{\eta}_h)_{\mathcal{C}}| \le 2\mu\|\boldsymbol{b}_h\|_0\|\boldsymbol{\eta}_h\|_0\,. \tag{4.26}$$

Now we choose the test functions $\boldsymbol{v}_h := \boldsymbol{u}_h - \boldsymbol{w}_h$ and $\boldsymbol{\eta}_h := \boldsymbol{a}_h - \boldsymbol{b}_h$, add the equations above, and obtain, by eventually applying (4.24) and (4.18),

$$
\begin{aligned}
&2\mu\|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h) + \boldsymbol{a}_h - \boldsymbol{b}_h\|_0^2 \\
&\quad \le (\boldsymbol{\epsilon}(\boldsymbol{u} - \boldsymbol{w}_h) - \boldsymbol{b}_h, \, \boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h))_{\mathcal{C}} - (\boldsymbol{\epsilon}(\boldsymbol{w}_h) + \boldsymbol{b}_h, \boldsymbol{a}_h - \boldsymbol{b}_h)_{\mathcal{C}} \\
&\qquad \le (\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \, \boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h))_0 + 2\mu\|\boldsymbol{b}_h\|_0\|\boldsymbol{a}_h - \boldsymbol{b}_h\|_0 \\
&\qquad\quad \le c_{20} h \|\boldsymbol{f}\|_0 \Big(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_0 + \|\boldsymbol{a}_h - \boldsymbol{b}_h\|_0 \Big).
\end{aligned} \tag{4.27}
$$

The strengthened Cauchy inequality (3.10) and Lemma 1(c) show, with (4.27) in the final step, that

$$
\begin{aligned}
&\sqrt{(1-c_5)/2}\Big(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_0^2 + \|\boldsymbol{a}_h - \boldsymbol{b}_h\|_0^2 \Big) \\
&\quad \le \|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h) + \boldsymbol{a}_h - \boldsymbol{b}_h\|_0^2 \\
&\quad \le c_{20}/2\mu \, h \|\boldsymbol{f}\|_0 \Big(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_0 + \|\boldsymbol{a}_h - \boldsymbol{b}_h\|_0 \Big).
\end{aligned}
$$

After dividing by the sum of the two norms from the last line and applying Young's inequality we arrive at

$$
\|\boldsymbol{\epsilon}(\boldsymbol{u}_h - \boldsymbol{w}_h)\|_0 + \|\boldsymbol{a}_h - \boldsymbol{b}_h\|_0 \le c_{21} h \|\boldsymbol{f}\|_0. \tag{4.28}
$$

This inequality together with (4.23) and the triangle inequality yields finally

$$
\|\boldsymbol{\epsilon}(\boldsymbol{u} - \boldsymbol{u}_h)\|_0 + \|\boldsymbol{a}_h\|_0 \le c_{22} h \|\boldsymbol{f}\|_0. \qquad \square \tag{4.29}
$$

Remark 6. We do have uniform convergence of the strains, but we do not have uniform convergence for the spherical part of the stresses. This corresponds to the fact that we do not get an inf-sup condition for the mixed method (3.8) that is independent of λ; cf. the discussion in Remark 4. In contrast to PEERS elements, we only have available trial functions for the stresses that are modifications of stresses of the form $\mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{v})$, $\boldsymbol{v} \in V^h$, that are derived from the assumed strains.

5 A posteriori error estimates

This section is devoted to computable upper error bounds for reliable error control. In the first step we justify our choice of norms on the spaces

$$
L := \{\boldsymbol{\sigma} \in L^2(\Omega; \mathbb{R}^{d\times d}_{sym}) : \int_\Omega \operatorname{tr}(\boldsymbol{\sigma}) dx = 0\} \quad \text{and} \quad V := H_0^1(\Omega)^d,
$$

namely, the L^2-norm on L and the H^1-norm in V. Recall that $\mathcal{C}$ is the fourth-order elasticity tensor which depends on the two material parameters μ and λ in (2.3).

The restriction on the traces in the definition of the space L is satisfied by all $\boldsymbol{\sigma} = \mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{u})$ whenever $\boldsymbol{u} \in V$. Indeed, from $\operatorname{tr}(\boldsymbol{\sigma}) = (2\mu + d\lambda) \operatorname{div} \boldsymbol{u}$ and $\boldsymbol{u} = 0$ on Γ it follows by the divergence theorem that

$$\int_\Omega \operatorname{tr}(\boldsymbol{\sigma})\ dx = (d\lambda + 2\mu) \int_\Omega \operatorname{div} \boldsymbol{u}\ dx = (d\lambda + 2\mu) \int_{\partial\Omega} \boldsymbol{u} \cdot \boldsymbol{n}\, ds = 0. \tag{5.1}$$

Theorem 3 *The operator $\mathcal{A} : L \times V \to (L \times V)'$ defined by*

$$\langle \mathcal{A}(\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v}) \rangle := (\mathcal{C}^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau})_0 - (\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\boldsymbol{v}))_0 - (\boldsymbol{\tau}, \boldsymbol{\epsilon}(\boldsymbol{u}))_0$$

for all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in L$ and $\boldsymbol{u}, \boldsymbol{v} \in V$, is bounded and bijective, and the operator norms of $\mathcal{A}$ and $\mathcal{A}^{-1}$ are λ-independent.

Remark 7. (a) The operator $\mathcal{A}$ in the theorem belongs to a bilinear form associated with the Hellinger-Reissner principle; compare with Problem E_h when recast in terms of stresses.

(b) The discrete situation involves $\boldsymbol{\sigma}_h := \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h)$, and $\mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{u}_h)$ belongs to L because of (5.1) while $\mathcal{C}\boldsymbol{a}_h$ belongs to L by assumption (as every component in (3.4) has integral mean zero over each element, hence over Ω).

(c) The theorem asserts that the chosen norms provide a λ-robust isomorphism and so motivates our choice of norms. An analogous result for energy norms (e.g., take the norm induced by the scalar product $(\mathcal{C}^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau})_0$ on L) is unknown.

Proof of Theorem 3. A proof is given for completeness. For this purpose we refer to the general theory of mixed formulations [4,6]. The continuity and inf-sup condition on $(\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\boldsymbol{u}))_0$ are well established, with λ-independent constants. The kernel of this bilinear form reads

$$Z := \{\boldsymbol{\sigma} \in L : \operatorname{div} \boldsymbol{\sigma} = 0\}$$

(where div is understood in the distributional sense). From the representation (2.4) we infer that the continuity bound of the bilinear form $(\mathcal{C}^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau})_0$ is λ-independent as well. Next, we need to check the ellipticity of the quadratic form $(\mathcal{C}^{-1}\boldsymbol{\tau}, \boldsymbol{\tau})_0$ only on the kernel Z. (It is no drawback that ellipticity on the larger set L holds only with a λ-dependent constant).

To verify uniform Z-ellipticity, we follow ideas from the proof of Lemma 4.1 in [7]. Given $\boldsymbol{\tau} \in Z \subset L$, we know that $\int_\Omega \operatorname{tr}(\boldsymbol{\tau}) dx = 0$. By Proposition IV.3.1 in [6] we have

$$\|\boldsymbol{\tau}\|_0 \leq c_{23}(\|\operatorname{dev}\boldsymbol{\tau}\|_0 + \|\operatorname{div}\boldsymbol{\tau}\|_0).$$

Here, $\operatorname{div}\boldsymbol{\tau} = 0$ holds. Since the deviatoric parts and the traces are L^2 orthogonal, it follows from $\operatorname{dev}\mathcal{C}^{-1}\boldsymbol{\tau} = \frac{1}{2\mu}\operatorname{dev}\boldsymbol{\tau}$ and $\operatorname{tr}(\mathcal{C}^{-1}\boldsymbol{\tau}) = \frac{1}{2\mu+d\lambda}\operatorname{tr}(\boldsymbol{\tau})$ that

$$\|\operatorname{dev}\boldsymbol{\tau}\|_0^2 = 2\mu\,(\operatorname{dev}\boldsymbol{\tau},\, \operatorname{dev}(\mathcal{C}^{-1}\boldsymbol{\tau}))_0 \leq 2\mu\,(\boldsymbol{\tau}, \mathcal{C}^{-1}\boldsymbol{\tau})_0\,.$$

By combining the last two inequalities we obtain

$$(\boldsymbol{\tau}, \mathcal{C}^{-1}\boldsymbol{\tau})_0 \geq \frac{1}{2\mu}\|\operatorname{dev}\boldsymbol{\tau}\|_0^2 \geq \frac{1}{2\mu c_{23}^2}\|\boldsymbol{\tau}\|_0^2\,. \quad \square$$

The assumption on the triangulation and elements in previous sections can be relaxed for our a posteriori error estimates. We suppose that $\boldsymbol{u}_h \in V^h \subset V$ and $\boldsymbol{a}_h \in \tilde{E}^h \subset L$ are $\mathcal{T}$-piecewise smooth functions (i.e., below, $\operatorname{div}_\mathcal{T}\boldsymbol{\sigma}_h \in L^2(\Omega)^2$, and $[\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_\mathcal{E}] \in L^2(\cup\mathcal{E})^2)$, where $\mathcal{T}$ is a regular triangulation of Ω into closed parallelograms. In terms of $\boldsymbol{\sigma}_h := \mathcal{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{a}_h)$, Eq. $(4.21)_1$ reads

$$\int_\Omega \boldsymbol{\sigma}_h : \boldsymbol{\epsilon}(\boldsymbol{v}_h)\, dx = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h\, dx \quad \text{for all } \boldsymbol{v}_h \in V^h = \mathcal{S}_0^1(\mathcal{T})$$

for the conforming $\mathcal{Q}_1$ finite element space $\mathcal{S}_0^1(\mathcal{T}) \subset H_0^1(\Omega)$ subordinated to the triangulation $\mathcal{T}$ and

$$\int_\Omega \operatorname{tr}(\boldsymbol{a}_h)\, dx = 0.$$

Together with symmetry of $\boldsymbol{a}_h$, this yields $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in L$.

To describe the a posteriori upper error bound of $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ recall that $\operatorname{div}_\mathcal{T}$ denotes the $\mathcal{T}$-piecewise divergence on Ω, and let $[\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_\mathcal{E}]$ denote the jump of $\boldsymbol{\sigma}_h$ on an edge $E \in \mathcal{E}$ with unit normal $\boldsymbol{n}_\mathcal{E}|_E = \boldsymbol{n}_E$. Let $h_\mathcal{T}$ and $h_\mathcal{E}$ denote the sizes of the meshes on Ω and of the edges on $\cup\mathcal{E}$, respectively.

Theorem 4 *There exists some $(\lambda, h_\mathcal{T}, h_\mathcal{E})$-independent constant $c_{24} > 0$ (that depends on Ω, the aspect ratio of the elements, and μ) with*

$$\begin{aligned} &c_{24}(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|\boldsymbol{u} - \boldsymbol{u}_h\|_1) \qquad (5.2)\\ &\leq \mu\|\boldsymbol{a}_h\|_0 + \|h_\mathcal{T}(\boldsymbol{f} + \operatorname{div}_\mathcal{T}\boldsymbol{\sigma}_h)\|_0 + \|h_\mathcal{E}^{1/2}[\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_\mathcal{E}]\|_{L^2(\cup\mathcal{E})}. \end{aligned}$$

Remark 8. (a) The reliable error estimate in Theorem 4 is efficient as well. Adopting arguments from [20] one can prove [21] (with λ-independent constant c_{25})

$$\|h_{\mathcal{T}}(\boldsymbol{f}+\operatorname{div}_{\mathcal{T}}\boldsymbol{\sigma}_h)\|_0+\|h_{\mathcal{E}}^{1/2}[\boldsymbol{\sigma}_h\cdot\boldsymbol{n}_{\mathcal{E}}]\|_{L^2(\cup\mathcal{E})}\le c_{25}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0+\text{ h.o.t.,} \tag{5.3}$$

where h.o.t. denotes higher order terms $\|h_{\mathcal{T}}(\boldsymbol{f}-\boldsymbol{f}_{\mathcal{T}})\|_0$ for the $\mathcal{T}$-piecewise constant integral means $f_{\mathcal{T}}$.

(b) The enhancement $\boldsymbol{a}_h$ of the strain can be understood as a residue of (2.3), $\boldsymbol{a}_h=\boldsymbol{\epsilon}(\boldsymbol{u}-\boldsymbol{u}_h)-\mathcal{C}^{-1}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)$ and

$$\|\boldsymbol{a}_h\|_0\le c_{26}\|\,\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\,\|_0+\|\,\boldsymbol{u}-\boldsymbol{u}_h\,\|_1. \tag{5.4}$$

According to (2.4), the constant c_{26} is independent of λ.

(c) Inhomogeneous Dirichlet data and Neumann conditions can be included as well. If Neumann conditions are present, we can omit the side restriction in L; cf. [7] for arguments in this direction.

(d) The description is restricted to parallelograms, but the theorem is valid for triangles or three dimensional elements as well. However, the enhanced strain method is degenerate for triangular elements; cf. [15].

The proof of Theorem 4 requires weak interpolation of Clément type.

Lemma 4 ([10]) *There exists an operator $J:V\to\mathcal{S}_0^1(\mathcal{T})^d$ such that, for all $\boldsymbol{v}\in V$,*

$$\|h_{\mathcal{T}}^{-1}(\boldsymbol{v}-J\boldsymbol{v})\|_{L^2(\Omega)}+\|h_{\mathcal{E}}^{-1/2}(\boldsymbol{v}-J\boldsymbol{v})\|_{L^2(\cup\mathcal{E})}\le c_{27}\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}. \quad\square \tag{5.5}$$

Proof of Theorem 4. The stability of $\mathcal{A}$ in Theorem 3 asserts that $\mathcal{A}$ satisfies an inf-sup condition. Since $\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\in L$ and $\boldsymbol{u}-\boldsymbol{u}_h\in V$, for some constant c_{28}, we find $\boldsymbol{\tau}\in L$ and $\boldsymbol{v}\in V$ with

$$\|\,\boldsymbol{\tau}\,\|_0+\|\,\boldsymbol{v}\,\|_1\le c_{28} \tag{5.6}$$

and

$$\begin{aligned}\|\,\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\,\|_0&+\|\,\boldsymbol{u}-\boldsymbol{u}_h\,\|_1\\&=\langle\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h),(\boldsymbol{\tau},\boldsymbol{v})\rangle\\&=(\mathcal{C}^{-1}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)-\boldsymbol{\epsilon}(\boldsymbol{u}-\boldsymbol{u}_h),\,\boldsymbol{\tau})_0-(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\,\boldsymbol{\epsilon}(\boldsymbol{v}))_0.\end{aligned}$$

Owing to $\boldsymbol{\sigma} = \mathcal{C}\boldsymbol{\epsilon}(\boldsymbol{u})$ and $\mathcal{C}^{-1}\boldsymbol{\sigma}_h - \boldsymbol{\epsilon}(\boldsymbol{u}_h) = \boldsymbol{a}_h$, the Galerkin orthogonality yields

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h,\, \boldsymbol{\epsilon}(\boldsymbol{v}_h))_0 = 0 \qquad \text{for } \boldsymbol{v}_h := J\boldsymbol{v}.$$

Elementwise integration by parts, and Cauchy's inequality yields

$$\begin{aligned}
&\langle \mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h), (\boldsymbol{\tau}, \boldsymbol{v})\rangle = -(\boldsymbol{a}_h, \boldsymbol{\tau})_0 - (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\epsilon}(\boldsymbol{v} - J\boldsymbol{v}))_0 \\
&\leq \| \boldsymbol{a}_h \|_0 \| \boldsymbol{\tau} \|_0 - (\boldsymbol{f} + \operatorname{div} \boldsymbol{\sigma}_h, \boldsymbol{v} - J\boldsymbol{v})_0 + \int_{\cup\mathcal{E}} [\boldsymbol{\sigma}_h \boldsymbol{n}_{\mathcal{E}}](\boldsymbol{v} - J\boldsymbol{v})\, ds \\
&\leq c_{28} \| \boldsymbol{a}_h \|_0 + \| h_{\mathcal{T}}^{-1}(\boldsymbol{v} - J\boldsymbol{v}) \|_0 \| h_{\mathcal{T}}(\boldsymbol{f} + \operatorname{div}_{\mathcal{T}} \boldsymbol{\sigma}_h) \|_0 \\
&\quad + \| h_{\mathcal{E}}^{-1/2}(\boldsymbol{v} - J\boldsymbol{v}) \|_{L^2(\cup\mathcal{E})} \| h_{\mathcal{E}}^{1/2} [\boldsymbol{\sigma}_h \boldsymbol{n}_{\mathcal{E}}] \|_{L^2(\cup\mathcal{E})}.
\end{aligned}$$

This inequality, Lemma 4, and (5.6) imply that

$$\begin{aligned}
&\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_0 + \| \boldsymbol{u} - \boldsymbol{u}_h \|_1 \\
&\quad \leq c_{28} \| \boldsymbol{a}_h \|_0 + c_{27} c_{28} \| h_{\mathcal{T}}(\boldsymbol{f} + \operatorname{div}_{\mathcal{T}} \boldsymbol{\sigma}_h) \|_0 \\
&\quad + c_{27} c_{28} \| h_{\mathcal{E}}^{1/2} [\boldsymbol{\sigma}_h \boldsymbol{n}_{\mathcal{E}}] \|_{L^2(\cup\mathcal{E})}. \qquad \square
\end{aligned}$$

Acknowledgements Part of the work of second author (CC) was performed while he was supported by the Max Planck Institute for Mathematics in the Sciences at Leipzig, Germany. The work of the third author (BDR) was supported by the National Research Foundation of South Africa. Moreover, the authors are indebted to the unknown referees for indicating an error in the first version of the manuscript.

References

1. Douglas N. Arnold, and L. Ridgway Scott and Michael Vogelius. *Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. Seris IV, **15**, 2, 169–192 (1988).
2. Manfred Bischoff, Ekkehard Ramm and Dietrich Braess. *A class of equivalent enhanced assumed strain and hybrid stress finite elements.* Comput. Mech., **22**, 6, 443-449 (1999).
3. Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer Verlag, New York, **15**, xii+294 (1994).
4. Dietrich Braess. *Finite Elements.* Cambridge University Press, Cambridge, xviii+352 (2001).
5. Dietrich Braess. *Enhanced assumed strain elements and locking in membrane problems.* Comput. Methods Appl. Mech. Engrg., **165**, 1-4, 155–174 (1998).
6. Franco Brezzi and Michel Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, New York x+350 (1991).
7. Carsten Carstensen and Georg Dolzmann. *A posteriori error estimates for mixed FEM in elasticity.* Numer. Math., **81**, 2, 187–209 (1998).

8. Carsten Carstensen, Georg Dolzmann, G., Stefan A. Funken and Dörte S. Helm. *Locking-free adaptive mixed finite element methods in linear elasticity.* Comput. Methods Appl. Mech. Engrg., **190**, 13–14, 1701–1718 (2000).
9. Carsten Carstensen and Stefan A. Funken. *Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods.* East-West J. Numer. Math., **8**, 3, 153–175 (2000).
10. Philippe Clément. *Approximation by finite element functions using local regularization.* Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér., **9**, R-2, 77–84 (1975).
11. Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer-Verlag, Berlin, x+374 (1986).
12. Claes Johnson and Juhani Pitkäranta. *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation.* Math. Comp., **46**, 173, 1–26 (1986).
13. C. Lovadina. *Analysis of strain-pressure finite element methods for the Stokes problem.* Numer. Methods Partial Differential Equations, **13**, 6, 717–730 (1997).
14. Jerrold E. Marsden and Thomas J.R. Hughes. *Mathematical Foundations of Elasticity.* Dover Publications Inc., New York, xviii+556 (1994).
15. B. Daya Reddy and J.C. Simo. *Stability and convergence of a class of enhanced strain methods.* SIAM J. Numer. Anal., **32**, 6, 1705–1728 (1995).
16. Stefanie Reese, P. Wriggers and B. Daya Reddy. *A new locking-free brick element technique for large deformation problems in elasticity.* Comput. & Structures, **75**, 3, 291–304 (2000).
17. J.C. Simo and F. Armero. *Geometrically nonlinear enhanced strain mixed methods and the method of incompatible modes.* Internat. J. Numer. Methods Engrg., **33**, 7, 1413–1449 (1992).
18. J.C. Simo, F. Armero and Robert L. Taylor. *Improved versions of assumed enhanced strain trilinear elements for 3D finite deformation problems.* Comput. Methods Appl. Mech. Engrg., **110**, 3-4, 359–386 (1993).
19. J.C. Simo and Rifai, M. S. (1990). *A class of mixed assumed strain methods and the method of incompatible modes.* Internat. J. Numer. Methods Engrg., **29**, 8, 1595–1638.
20. Rüdiger Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner (1996).
21. Rüdiger Verfürth. *A review of a posteriori error estimation techniques for elasticity problems.* Comput. Methods Appl. Mech. Engrg., **176**, 1-4, 419–440 (1999).
22. Michael Vogelius. *An analysis of the p-version of the finite element method for nearly incompressible materials. Uniformly valid, optimal error estimates.* Numer. Math., **41**, 1, 39–53 (1983).
23. Seung T. Yeo and Byung C. Lee. *Equivalence between enhanced assumed strain method and assumed stress hybrid method based on the Hellinger-Reissner principle.* Internat. J. Numer. Methods Engrg., **39**, 18, 3083–3099 (1996).