An Adaptive Least Squares Mixed Finite Element Method for the Stress-Displacement Formulation of Linear Elasticity

Zhiqiang Cai,[1] **Johannes Korsawe,**[2] **Gerhard Starke**[2]

[1]*Department of Mathematics, Purdue University, 1395 Mathematical Sciences Building, West Lafayette, Indiana 47907-1395*

[2]*Institut für Angewandte Mathematik, Universität Hannover, Welfengarten 1, 30167 Hannover, Germany*

A least-squares mixed finite element method for linear elasticity, based on a stress-displacement formulation, is investigated in terms of computational efficiency. For the stress approximation quadratic Raviart-Thomas elements are used and these are coupled with the quadratic nonconforming finite element spaces of Fortin and Soulie for approximating the displacement. The local evaluation of the least-squares functional serves as an *a posteriori* error estimator to be used in an adaptive refinement algorithm. We present computational results for a benchmark test problem of planar elasticity including nearly incompressible material parameters in order to verify the effectiveness of our adaptive strategy. For comparison, conforming quadratic finite elements are also used for the displacement approximation showing convergence orders similar to the nonconforming case, which are, however, not independent of the Lamé parameters. © 2004 Wiley Periodicals, Inc. Numer Methods Partial Differential Eq 21: 132–148, 2005

Keywords: linear elasticity; stress displacement

I. INTRODUCTION

The purpose of this article is a detailed investigation of the computational performance of a least-squares mixed finite element approach, based on a stress-displacement formulation, for linear elasticity. The least-squares formulation treated in this article is closely related to one of the approaches investigated in [1], where the functional contains an additional least-squares term enforcing symmetry of the stress tensor. It will be shown below that this additional term is actually not needed. In other words, the symmetry of the stress tensor is enforced weakly by the

variational formulation. Our emphasis in this paper is on the performance of an adaptive refinement strategy based on the *a posteriori* error estimator inherent in the least-squares formulation by the local evaluation of the functional. The numerical results in [1] were of a preliminary nature involving a simple model problem with pure displacement boundary conditions, which was discretized with lowest-order (linear) Raviart-Thomas elements for the stress and (linear nonconforming) Crouzeix-Raviart elements for the displacement. In this article, we show the wider applicability of our method by numerical tests for a benchmark problem for planar linear elasticity taken from [2]. This requires the use of the quadratic nonconforming finite elements by Fortin-Soulie for the displacement approximation, which are coupled with quadratic Raviart-Thomas elements in our computations. To this end, the coercivity and approximation properties shown in [1] for the linear case are extended to higher-order nonconforming elements here. Most importantly, adaptive refinement based on the local evaluation of the least-squares functional leads to an efficient discretization method with approximation of optimal order.

Finite element methods of least-squares type have been the object of many studies recently (see, e.g., the survey [3] and the monograph [4]). Least-squares finite element methods have also been applied to first-order system formulations of linear elasticity, for example, in [5] where displacement gradients are used as additional degrees of freedom. In fact, mixed finite element approaches for elasticity problems, approximating simultaneously the displacement field and the stress tensor as independent variables, have a much longer history. The most popular of these methods use H(div)-conforming finite elements for the stress tensor combined with appropriate spaces for the displacement and an additional Lagrange multiplier for the antisymmetric part of the strain tensor. In this context, the PEERS element [6] is based on an extension of the Raviart-Thomas spaces, while the approach in [7] rests upon a suitable extension of the Brezzi-Douglas-Marini elements. Error estimators for such mixed approaches to linear elasticity have been developed in [8] and [9]. A detailed comparison of our least-squares finite element method with the other approaches mentioned above is beyond the scope of this article. In fact, such a comparison would be sensitive with respect to the choice of norm in which the different variables are to be approximated. Moreover, a fair comparison would also involve the use of the most efficient adaptive refinement strategies available for all competing methods. Instead of a full comparison of numerical results we comment on the expected performance of different finite element methods for linear elasticity in comparison to our least-squares approach at the end of Section 5.

The crucial ingredient of our least-squares formulation is the proper weighting of the stress-strain relation in comparison to the continuity equation. The use of the element-wise evaluation of the resulting least-squares functional as an *a posteriori* error estimator is emphasized in this article. The resulting adaptive refinement strategy is shown to be effective, uniformly in the incompressible regime, for a benchmark test problem. The approximation spaces to be used in the actual implementation of our least-squares formulation are quadratic Raviart-Thomas elements for the stress combined with the quadratic nonconforming elements by Fortin and Soulie [10] for the displacement. The use of standard quadratic conforming finite element spaces for the displacement components is also investigated. The obtained convergence rates are comparable to those for the nonconforming elements but deteriorate in the incompressible limit. Different least-squares approaches for the stress-displacement formulation were proposed in [11]. They differ from the one considered here mainly in the weighting of the individual terms in the differential system. Both approaches possess uniform approximation properties in the incompressible limit but with respect to different norms. The method studied in [11] is constructed in order to get uniform convergence with respect to a product norm

representing strain. In contrast, the approach in [1] and the one under consideration here achieve uniform convergence with respect to the energy norms for the stress and displacement variables.

An outline of the article is as follows. The least-squares formulation of the linear elasticity model based on stress and displacement is described in Section 2. This includes the coercivity properties of the least-squares variational formulation. Appropriate spaces for the finite element approximation and a generalization of the coercivity shown in Section 2 to the nonconforming case is discussed in Section 3. Section 4 presents the equivalence of the functional to an appropriate error norm and its consequences for *a posteriori* error estimation via local evaluation of the least-squares functional. Finally, computational results for a benchmark test problem are shown and commented in Section 5 along with some remarks on the performance relative to other finite element approaches to linear elasticity.

II. A LEAST-SQUARES FORMULATION OF LINEAR ELASTICITY

We start from the equations of linear elasticity in the form

$$\text{div } \boldsymbol{\sigma} = \mathbf{0} \qquad \text{in } \Omega,$$

$$\boldsymbol{\sigma} - \mathscr{C}\varepsilon(\mathbf{u}) = \mathbf{0} \qquad \text{in } \Omega,$$

$$\mathbf{u} = \mathbf{0} \qquad \text{on } \Gamma_D,$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{g} \qquad \text{on } \Gamma_N \qquad\qquad (2.1)$$

in a bounded domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$ partitioned into Γ_D and Γ_N such that $\bar{\Gamma}_D \cup \bar{\Gamma}_N = \partial\Omega$. For simplicity, we assume that $\Gamma_D \neq \varnothing$ and $\Gamma_N \neq \varnothing$. Although all the methods presented in this article can be extended to the three-dimensional case, we restrict our exposition to $d = 2$. We introduce the Sobolev spaces

$$H(\text{div}, \Omega) = \{\mathbf{s} \in L^2(\Omega)^2 : \text{div } \mathbf{s} \in L^2(\Omega)\},$$

$$H^1(\Omega) = \{p \in L^2(\Omega) : \nabla p \in L^2(\Omega)^2\}.$$

Solutions of (2.1) for $\boldsymbol{\sigma} : \Omega \to \mathbb{R}^{2\times2}$ (stress) and $\mathbf{u} : \Omega \to \mathbb{R}^2$ (displacement) are then sought in $H(\text{div}, \Omega)^2$ and $H^1(\Omega)^2$, respectively. Note that div $\boldsymbol{\tau}$ means row-wise application of the divergence operator. Similarly, $\nabla\mathbf{u}$ contains the gradient vectors of the components of $\mathbf{u}$ in each row. In (2.1), $\mathscr{C}$ describes the linear mapping from strains to stresses,

$$\mathscr{C}\varepsilon = 2\mu\varepsilon + \lambda(\text{tr }\varepsilon)\mathbf{I} = \begin{pmatrix} (2\mu + \lambda)\varepsilon_{11} + \lambda\varepsilon_{22} & 2\mu\varepsilon_{12} \\ 2\mu\varepsilon_{21} & \lambda\varepsilon_{11} + (2\mu + \lambda)\varepsilon_{22} \end{pmatrix}, \qquad (2.2)$$

$$\text{and} \quad \varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T) = \begin{pmatrix} \partial_1 u_1 & (\partial_2 u_1 + \partial_1 u_2)/2 \\ (\partial_2 u_1 + \partial_1 u_2)/2 & \partial_2 u_2 \end{pmatrix} \qquad (2.3)$$

denotes the linear strain tensor. The parameters λ and μ in (2.2) are the well-known Lamé constants associated with the material (cf. [12, Sect. VI.1]). We will also make use of the inverse mapping $\mathscr{C}^{-1}$ from stresses to strains in the sequel,

$$\mathcal{C}^{-1}\boldsymbol{\sigma} = \frac{1}{2\mu}\,\boldsymbol{\sigma} - \frac{\lambda}{4\mu(\lambda+\mu)}\,(\text{tr }\boldsymbol{\sigma})\mathbf{I}$$

$$= \frac{1}{2\mu}\begin{pmatrix} \dfrac{2\mu+\lambda}{2(\lambda+\mu)}\,\sigma_{11} - \dfrac{\lambda}{2(\lambda+\mu)}\,\sigma_{22} & \sigma_{12} \\[2ex] \sigma_{21} & -\dfrac{\lambda}{2(\lambda+\mu)}\,\sigma_{11} + \dfrac{2\mu+\lambda}{2(\lambda+\mu)}\,\sigma_{22} \end{pmatrix}. \quad (2.4)$$

By $\boldsymbol{\sigma}^N \in H(\text{div},\Omega)^2$ we denote a suitable extension of the boundary values, i.e., such that $\boldsymbol{\sigma}^N \cdot \mathbf{n} = \mathbf{g}$ on Γ_N. Our aim is to find $\boldsymbol{\sigma} = \boldsymbol{\sigma}^N + \hat{\boldsymbol{\sigma}}$ and $\mathbf{u}$ with

$$\hat{\boldsymbol{\sigma}} \in H_{\Gamma_N}(\text{div},\Omega)^2 = \{\boldsymbol{\tau} \in H(\text{div},\Omega)^2 : \boldsymbol{\tau}\cdot\mathbf{n} = \mathbf{0} \text{ on } \Gamma_N\},$$

$$\mathbf{u} \in H^1_{\Gamma_D}(\Omega)^2 = \{\mathbf{v} \in H^1(\Omega)^2 : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D\}$$

such that (2.1) is satisfied. To this end, we introduce the least-squares functional

$$\mathcal{F}(\boldsymbol{\sigma},\mathbf{u}) = \|\text{div }\boldsymbol{\sigma}\|^2_{0,\Omega} + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\sigma} - \mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{u})\|^2_{0,\Omega}. \quad (2.5)$$

The first-order system (2.1) may be solved by minimizing $\mathcal{F}(\boldsymbol{\sigma},\mathbf{u})$ among all $(\boldsymbol{\sigma},\mathbf{u}) \in (\boldsymbol{\sigma}^N + H_{\Gamma_N}(\text{div},\Omega)^2) \times H^1_{\Gamma_D}(\Omega)^2$. Associated with this quadratic functional is the bilinear form

$$\mathcal{B}(\boldsymbol{\sigma},\mathbf{u};\boldsymbol{\tau},\mathbf{v}) = (\text{div }\boldsymbol{\sigma}, \text{div }\boldsymbol{\tau})_{0,\Omega} + \mu(\mathcal{C}^{-1/2}\boldsymbol{\sigma} - \mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{u}), \mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v}))_{0,\Omega} \quad (2.6)$$

for $(\boldsymbol{\sigma},\mathbf{u}), (\boldsymbol{\tau},\mathbf{v}) \in H(\text{div},\Omega)^2 \times H^1(\Omega)^2$ and the linear functional

$$\mathcal{L}(\boldsymbol{\tau},\mathbf{v}) = -(\text{div }\boldsymbol{\sigma}^N, \text{div }\boldsymbol{\tau})_{0,\Omega} - \mu(\boldsymbol{\sigma}^N, \mathcal{C}^{-1}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\mathbf{v}))_{0,\Omega} \quad (2.7)$$

for $(\boldsymbol{\tau},\mathbf{v}) \in H(\text{div},\Omega)^2 \times H^1(\Omega)^2$. The minimization of the quadratic functional in (2.5) is then equivalent to the linear variational problem of finding $\boldsymbol{\sigma} = \boldsymbol{\sigma}^N + \hat{\boldsymbol{\sigma}}$ with $\hat{\boldsymbol{\sigma}} \in H_{\Gamma_N}(\text{div},\Omega)^2$ and $\mathbf{u} \in H^1_{\Gamma_D}(\Omega)^2$ such that

$$\mathcal{B}(\hat{\boldsymbol{\sigma}},\mathbf{u};\boldsymbol{\tau},\mathbf{v}) = \mathcal{L}(\boldsymbol{\tau},\mathbf{v}) \qquad \text{for all } (\boldsymbol{\tau},\mathbf{v}) \in H_{\Gamma_N}(\text{div},\Omega)^2 \times H^1_{\Gamma_D}(\Omega)^2. \quad (2.8)$$

On the product space $H_{\Gamma_N}(\text{div},\Omega)^2 \times H^1_{\Gamma_D}(\Omega)^2$, we define the norm $\|\|(\cdot,\cdot)\|\|$ by

$$\|\|(\boldsymbol{\tau},\mathbf{v})\|\| = (\|\text{div }\boldsymbol{\tau}\|^2_{0,\Omega} + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|^2_{0,\Omega} + \mu\|\mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|^2_{0,\Omega})^{1/2}. \quad (2.9)$$

In order to get some insight into the properties of this norm, we split $\mathcal{C}^{-1}$ into its deviatoric and volumetric parts as

$$\mathcal{C}^{-1}\boldsymbol{\tau} = \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{1}{2}(\text{tr }\boldsymbol{\tau})\mathbf{I}\right) + \frac{1}{4(\lambda+\mu)}(\text{tr }\boldsymbol{\tau})\mathbf{I} = \frac{1}{2\mu}\text{ dev }\boldsymbol{\tau} + \frac{1}{2(\lambda+\mu)}\text{ vol }\boldsymbol{\tau},$$

which implies

$$\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 = \frac{1}{2\mu}\|\text{dev } \boldsymbol{\tau}\|_{0,\Omega}^2 + \frac{1}{4(\lambda + \mu)}\|\text{tr } \boldsymbol{\tau}\|_{0,\Omega}^2. \tag{2.10}$$

This means that for large λ the deviatoric part of the stress dominates the norm. Based on a similar decomposition for the displacement,

$$\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{v}) = 2\mu(\boldsymbol{\varepsilon}(\mathbf{v}) - \tfrac{1}{2}(\text{div } \mathbf{v})\mathbf{I}) + (\lambda + \mu)(\text{div } \mathbf{v})\mathbf{I},$$

the corresponding part of the norm satisfies

$$\|\mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 = 2\mu\|\boldsymbol{\varepsilon}(\mathbf{v}) - \tfrac{1}{2}(\text{div } \mathbf{v})\mathbf{I}\|_{0,\Omega}^2 + (\lambda + \mu)\|\text{div } \mathbf{v}\|_{0,\Omega}^2 = 2\mu\|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 + \lambda\|\text{div } \mathbf{v}\|_{0,\Omega}^2,$$
$$\tag{2.11}$$

which means that it is dominated by the volumetric part of the displacement.

Under the above assumptions, it is well known that the following Korn's inequality holds for all $\mathbf{v} \in H^1_{\Gamma_D}(\Omega)^2$ (cf. Braess [12, Section VI.3]):

$$\|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla\mathbf{v}\|_{0,\Omega}^2 \leq C_K\|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2$$

with a constant C_K. With (2.11) this implies

$$2\mu(\|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla\mathbf{v}\|_{0,\Omega}^2) \leq 2\mu C_K\|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 \leq C_K\|\mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2. \tag{2.12}$$

Clearly, since

$$\|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 = \|\partial_1 v_1\|_{0,\Omega}^2 + 2\|\tfrac{1}{2}(\partial_1 v_2 + \partial_2 v_1)\|_{0,\Omega}^2 + \|\partial_2 v_2\|_{0,\Omega}^2 \leq \|\partial_1 v_1\|_{0,\Omega}^2 + \|\partial_1 v_2\|_{0,\Omega}^2$$
$$+ \|\partial_2 v_1\|_{0,\Omega}^2 + \|\partial_2 v_2\|_{0,\Omega}^2 = \|\nabla\mathbf{v}\|_{0,\Omega}^2,$$

we have $C_K \geq 1$. We remark that (2.10) and (2.11) combined with Korn's inequality (2.12) imply that $\||(\cdot, \cdot)\||$, defined in (2.9), is indeed a norm on $H_{\Gamma_N}(\text{div}, \Omega)^2 \times H^1_{\Gamma_D}(\Omega)^2$.

Our aim is to get approximation results which are uniform in the Lamé parameters λ and μ, in particular, for $\lambda \rightarrow \infty$. Since, in practical applications, μ may also be quite large the functional and norm are also scaled by this parameter. Alternatively, one could assume that μ is on the order of one and rescale the variables and right-hand side accordingly. The well-posedness of the variational problem (2.8) follows from continuity and coercivity of the bilinear form (2.6) in $H_{\Gamma_N}(\text{div}, \Omega)^2 \times H^1_{\Gamma_D}(\Omega)^2$.

Theorem 2.1. *The bilinear form $\mathcal{B}(\cdot, \cdot; \cdot, \cdot)$ is continuous and coercive, uniformly in μ and λ, with respect to $\||(\cdot, \cdot)\||$. In other words, there exist positive constants α and β, independent of μ and λ, such that*

$$\mathcal{B}(\boldsymbol{\sigma}, \mathbf{u}; \boldsymbol{\tau}, \mathbf{v}) \leq \beta \||(\boldsymbol{\sigma}, \mathbf{u})\|| \; \||(\boldsymbol{\tau}, \mathbf{v})\||, \tag{2.13}$$

$$\mathcal{B}(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) \geq \alpha \||(\boldsymbol{\tau}, \mathbf{v})\||^2 \tag{2.14}$$

holds for all $(\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v}) \in H_{\Gamma_N}(\text{div}, \Omega)^2 \times H^1_{\Gamma_D}(\Omega)^2$.

Proof.   (i) For the upper bound we have

$$\mathcal{B}(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) = \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 \le \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + 2\mu(\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2$$
$$+ \|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2) \le 2\, \|\|\, (\boldsymbol{\tau}, \mathbf{v})\, \|\|^{\,2}.$$

Since the bilinear form is symmetric, this is sufficient for the upper bound in Theorem 2.1.

(ii) For the proof of the lower bound we will repeatedly use the decomposition of the stress tensor into its symmetric and antisymmetric part,

$$\boldsymbol{\tau} = \mathrm{sy}\,\boldsymbol{\tau} + \mathrm{as}\,\boldsymbol{\tau} \qquad \text{with sy } \boldsymbol{\tau} = \frac{\boldsymbol{\tau} + \boldsymbol{\tau}^T}{2}, \qquad \text{as } \boldsymbol{\tau} = \frac{\boldsymbol{\tau} - \boldsymbol{\tau}^T}{2}.$$

Obviously,

$$(\mathrm{sy}\,\boldsymbol{\tau}, \mathrm{as}\,\boldsymbol{\tau})_{0,\Omega} = \tfrac{1}{4}((\boldsymbol{\tau}, \boldsymbol{\tau})_{0,\Omega} - (\boldsymbol{\tau}^T, \boldsymbol{\tau}^T)_{0,\Omega} + (\boldsymbol{\tau}^T, \boldsymbol{\tau})_{0,\Omega} - (\boldsymbol{\tau}, \boldsymbol{\tau}^T)_{0,\Omega}) = 0,$$

since the first and second term and the third and fourth term, respectively, cancel, i.e., the decomposition is orthogonal. Moreover,

$$\|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 \ge \|\mathrm{as}(\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v}))\|_{0,\Omega}^2 = \|\mathrm{as}(\mathcal{C}^{-1/2}\boldsymbol{\tau})\|_{0,\Omega}^2 = \frac{1}{2\mu}\|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 \quad (2.15)$$

holds due to the above orthogonality and since $\mathcal{C}^{1/2}\varepsilon(\mathbf{v})$ is symmetric. Again, the orthogonal decomposition into symmetric and antisymmetric part gives

$$(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} = (\mathrm{sy}\,\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} + (\mathrm{as}\,\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} = (\mathrm{sy}\,\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega}$$
$$= (\mathrm{sy}\,\boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega} = (\boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega} - (\mathrm{as}\,\boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega} = -(\mathrm{div}\,\boldsymbol{\tau}, \mathbf{v})_{0,\Omega} - (\mathrm{as}\,\boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega}. \quad (2.16)$$

This identity was also crucial in the ellipticity proof in [1]. We may now use (2.15) in order to get

$$\mathcal{B}(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) = \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 \ge \tfrac{1}{3}(\|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2$$
$$+ \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2) = \tfrac{1}{3}(\|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega}).$$
$$(2.17)$$

With constants $\gamma \in (0, 1)$ and $\delta > 0$ which will be fixed below and using (2.16) the term in brackets may be further bounded from below as

$$\|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} = \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2$$
$$+ \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu\gamma(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} + \mu(1 - \gamma)(\|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2$$
$$- \|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 - \|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2) \ge \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2$$
$$+ 2\mu\gamma(\mathrm{div}\,\boldsymbol{\tau}, \mathbf{v})_{0,\Omega} + 2\mu\gamma(\mathrm{as}\,\boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega} \ge \|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2$$
$$+ \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - \frac{\mu\gamma}{\delta}\|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - \mu\gamma\delta\|\mathbf{v}\|_{0,\Omega}^2 - \|\mathrm{as}\,\boldsymbol{\tau}\|_{0,\Omega}^2 - \mu^2\gamma^2\|\nabla\mathbf{v}\|_{0,\Omega}^2 = \left(1 - \frac{\mu\gamma}{\delta}\right)\|\mathrm{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2$$
$$+ \mu\gamma\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathcal{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 - \mu\gamma\delta\|\mathbf{v}\|_{0,\Omega}^2 - \mu^2\gamma^2\|\nabla\mathbf{v}\|_{0,\Omega}^2.$$
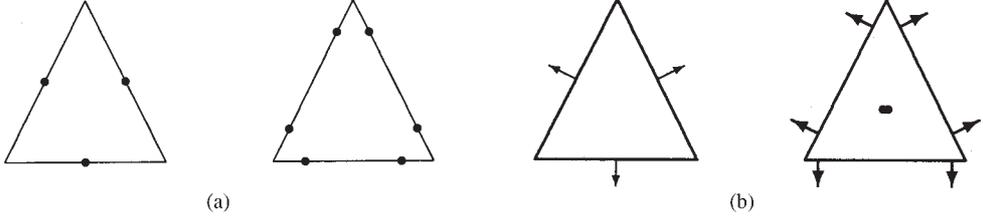
FIG. 1.   (a) Linear and quadratic elements on triangles. (b) Degrees of freedom for the Raviart-Thomas spaces.

Using Korn's inequality (2.12) and setting $\delta = 2\mu\gamma$, we obtain

$$\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\text{as } \boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} \geq \tfrac{1}{2}\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2$$

$$+ \mu\gamma\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 - 2\mu^2\gamma^2(\|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla\mathbf{v}\|_{0,\Omega}^2) \geq \tfrac{1}{2}\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\gamma\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2$$

$$+ \mu\gamma\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 - \mu\gamma^2 C_K\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2.$$

Finally, choosing $\gamma = 1/(2C_K)$ (which satisfies $\gamma \in (0, 1)$ since $C_K \geq 1$) leads to

$$\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 + \|\text{as } \boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu(\boldsymbol{\tau}, \varepsilon(\mathbf{v}))_{0,\Omega} \geq \frac{1}{2}\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2$$

$$+ \frac{\mu}{2C_K}\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \frac{\mu}{4C_K}\|\mathscr{C}^{1/2}\varepsilon(\mathbf{v})\|_{0,\Omega}^2 \geq \frac{1}{4C_K} \left\|\left\|(\boldsymbol{\tau}, \mathbf{v})\right\|\right\|^2.$$

Combined with (2.17), we obtain

$$\mathscr{B}(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) \geq \frac{1}{12C_K} \left\|\left\|(\boldsymbol{\tau}, \mathbf{v})\right\|\right\|^2. \qquad \blacksquare$$

III. FINITE ELEMENT APPROXIMATION

In principle, the least-squares mixed finite element approach simply consists of minimizing (2.5) in finite-dimensional subspaces $\mathbf{X}_h \subset H_{\Gamma_N}(\text{div}, \Omega)^2$ and $\mathbf{V}_h \subset H_{\Gamma_D}^1(\Omega)^2$. Suitable spaces are based on a triangulation $\mathscr{T}_h$ of Ω and consist of piecewise polynomials with sufficient continuity conditions. Additionally, the choice of spaces is restricted by the necessity to have approximation properties with respect to the $\|\|(\cdot, \cdot)\|\|$ norm uniformly in λ (cf. [13, Sect. 11.4], [14, Sect. VI.3]).

In particular, this complicates the displacement approximation since standard piecewise linear or piecewise quadratic finite elements do not yield approximation properties with respect to $\|\|(\cdot, \cdot)\|\|$ which are uniform in the incompressible limit, i.e., as $\lambda \rightarrow \infty$. Note that the part of the $\|\|(\cdot, \cdot)\|\|$ coincides with the energy norm for the standard displacement formulation and the problem of uniform finite element approximation is therefore identical. Uniform approximation properties can be accomplished with nonconforming finite element spaces where $\mathbf{V}_h \not\subset H_{\Gamma_D}^1(\Omega)^2$ and continuity is only enforced at the Gauss points on the common edge between two elements. This is illustrated for linear and quadratic elements on triangles in Fig. 1(a).

In the linear case, these are the well-known Crouzeix-Raviart elements which can be implemented in a straightforward way. In the quadratic case, a convenient implementation was suggested by Fortin and Soulie [10]. Displacement boundary conditions are also enforced weakly, i.e., $\mathbf{v}_h$ vanishes at the Gauss points on the edges $E \subseteq \Gamma_D$ of the triangulation. Since $\mathbf{V}_h \not\subseteq H^1_{\Gamma_D}(\Omega)^2$, $\varepsilon(\mathbf{v}_h) \notin L^2(\Omega)^{2 \times 2}$ and the least-squares functional (2.5) needs to be modified to the discrete functional

$$\mathscr{F}_h(\boldsymbol{\sigma}, \mathbf{u}) = \|\operatorname{div} \boldsymbol{\sigma}\|^2_{0,\Omega} + \mu \sum_{T \in \mathscr{T}_h} \|\mathscr{C}^{-1/2} \boldsymbol{\sigma} - \mathscr{C}^{1/2} \varepsilon(\mathbf{u})\|^2_{0,T}. \tag{3.1}$$

For the stress approximation the inclusion $\mathbf{X}_h \subset H_{\Gamma_N}(\operatorname{div}, \Omega)^2$ can actually be achieved by using Raviart-Thomas elements for each row of the stress tensor. The Raviart-Thomas finite element spaces are given by piecewise polynomials of the form

$$\mathbf{s}_{h|T} = \begin{pmatrix} p^{(I)}_{k-1} \\ p^{(II)}_{k-1} \end{pmatrix} + \mathbf{x} p^{(III)}_{k-1}$$

on each triangle $T \in \mathscr{T}_h$, where $p^{(I)}_{k-1}$, $p^{(II)}_{k-1}$, and $p^{(III)}_{k-1}$ denote polynomials of degree $k - 1$. For $k = 1$, this implies

$$\mathbf{s}_{h|T} = \begin{pmatrix} \alpha_T \\ \beta_T \end{pmatrix} + \gamma_T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and for $k = 2$,

$$\mathbf{s}_{h|T} = \begin{pmatrix} \alpha_T + \beta_T x_1 + \gamma_T x_2 \\ \delta_T + \rho_T x_1 + \sigma_T x_2 \end{pmatrix} + (\omega_T x_1 + \eta_T x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The Raviart-Thomas space of degree k is also characterized as those polynomials of degree k (component-wise) with the property that the trace of $\mathbf{s}_h \cdot \mathbf{n}$ on each edge is a polynomial of degree $k - 1$. For this reason the notation RT_{k-1} is commonly used for the Raviart-Thomas space consisting of piecewise polynomials of degree k. A convenient choice of basis functions to represent the degrees of freedom for the Raviart-Thomas spaces is indicated in Fig. 1(b) (cf. [14, Sect. III.3]).

Minimizing the functional (3.1) among all $\boldsymbol{\sigma}_h = \boldsymbol{\sigma}^N + \hat{\boldsymbol{\sigma}}_h$ with $\hat{\boldsymbol{\sigma}}_h \in \mathbf{X}_h$ and among all $\mathbf{u}_h \in \mathbf{V}_h$ is then equivalent to the following variational problem: Find $\hat{\boldsymbol{\sigma}}_h \in \mathbf{X}_h$ and $\mathbf{u}_h \in \mathbf{V}_h$ such that

$$\mathscr{B}_h(\hat{\boldsymbol{\sigma}}_h, \mathbf{u}_h; \boldsymbol{\tau}, \mathbf{v}) = \mathscr{L}_h(\boldsymbol{\tau}, \mathbf{v}) \tag{3.2}$$

holds for all $\boldsymbol{\tau} \in \mathbf{X}_h$ and $\mathbf{v} \in \mathbf{V}_h$. The discrete bilinear form $\mathscr{B}_h(\cdot, \cdot; \cdot, \cdot)$ and linear functional $\mathscr{L}_h(\cdot, \cdot)$ are defined as follows:

$$\mathscr{B}_h(\boldsymbol{\sigma}, \mathbf{u}; \boldsymbol{\tau}, \mathbf{v}) = (\operatorname{div} \boldsymbol{\sigma}, \operatorname{div} \boldsymbol{\tau})_{0,\Omega} + \mu \sum_{T \in \mathscr{T}_h} (\mathscr{C}^{-1/2} \boldsymbol{\sigma} - \mathscr{C}^{1/2} \varepsilon(\mathbf{u}), \mathscr{C}^{-1/2} \boldsymbol{\tau} - \mathscr{C}^{1/2} \varepsilon(\mathbf{v}))_{0,T}, \quad (3.3)$$

$$\mathscr{L}_h(\boldsymbol{\tau}, \mathbf{v}) = -(\operatorname{div} \boldsymbol{\sigma}^N, \operatorname{div} \boldsymbol{\tau})_{0,\Omega} - \mu \sum_{T \in \mathscr{T}_h} (\boldsymbol{\sigma}^N, \mathscr{C}^{-1}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T}. \tag{3.4}$$

With respect to the discrete analogue of the norm (2.9),

$$\||(\boldsymbol{\tau}, \mathbf{v})\||_h = \left(\|\operatorname{div} \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu \|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \mu \sum_{T \in \mathscr{T}_h} \|\mathscr{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \right)^{1/2}, \tag{3.5}$$

we will show a similar result to Theorem 2.1 below. In order to ensure that $\||(\cdot, \cdot)\||_h$ defines a norm, this requires a discrete analogue of Korn's inequality of the type

$$\|\mathbf{v}\|_{0,\Omega}^2 + \sum_{T \in \mathscr{T}_h} \|\nabla \mathbf{v}\|_{0,T}^2 \leq C_K' \sum_{T \in \mathscr{T}_h} \|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2$$

for all $\mathbf{v} \in \mathbf{V}_h$ with a constant C_K'. Note that in analogy to (2.12) the discrete Korn's inequality implies that

$$2\mu \left(\|\mathbf{v}\|_{0,\Omega}^2 + \sum_{T \in \mathscr{T}_h} \|\nabla \mathbf{v}\|_{0,T}^2 \right) \leq C_K' \sum_{T \in \mathscr{T}_h} \|\mathscr{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \tag{3.6}$$

holds and that $C_K' \geq 1$ for the same reason as for the standard Korn's inequality.

Unfortunately, the discrete Korn's inequality is not valid, in general, if $\Gamma_N \neq \emptyset$, for the Crouzeix-Raviart elements (cf. [14, Section VI.3]). For the quadratic nonconforming finite element space, a discrete Korn's inequality of the above form was shown to hold in [15].

Theorem 3.1. *Let $\mathbf{V}_h$ be a nonconforming finite element space of degree k which satisfies the discrete Korn's inequality and let $\mathbf{X}_h$ be the Raviart-Thomas space of degree $k' \leq k$. Then, there exist positive constants α and β, which are independent of μ and λ, such that*

$$\mathscr{B}_h(\boldsymbol{\sigma}, \mathbf{u}; \boldsymbol{\tau}, \mathbf{v}) \leq \beta \||(\boldsymbol{\sigma}, \mathbf{u})\||_h \||(\boldsymbol{\tau}, \mathbf{v})\||_h, \tag{3.7}$$

$$\mathscr{B}_h(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) \geq \alpha \||(\boldsymbol{\tau}, \mathbf{v})\||_h^2 \tag{3.8}$$

holds for all $(\boldsymbol{\sigma}, \mathbf{u}) \in H_{\Gamma_N}(\operatorname{div}, \Omega)^2 \times (H_{\Gamma_D}^1(\Omega)^2 + \mathbf{V}_h)$ and for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X}_h \times \mathbf{V}_h$.

Proof. The proof is similar to the one for Theorem 2.1 if $\mathscr{B}(\cdot, \cdot; \cdot, \cdot)$ is replaced by $\mathscr{B}_h(\cdot, \cdot; \cdot, \cdot)$ and $\||(\cdot, \cdot)\||$ by $\||(\cdot, \cdot)\||_h$. The upper bound follows from

$$\mathscr{B}_h(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) = \|\operatorname{div} \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu \sum_{T \in \mathscr{T}_h} \|\mathscr{C}^{-1/2}\boldsymbol{\tau} - \mathscr{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \leq \|\operatorname{div} \boldsymbol{\tau}\|_{0,\Omega}^2$$

$$+ 2\mu \left(\|\mathscr{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \sum_{T \in \mathscr{T}_h} \|\mathscr{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \right) \leq 2 \||(\boldsymbol{\tau}, \mathbf{v})\||_h^2,$$

which clearly holds for all $\boldsymbol{\tau} \in H_{\Gamma_N}(\operatorname{div}, \Omega)^2$ and for all $\mathbf{v} \in H_{\Gamma_D}^1(\Omega)^2 + \mathbf{V}_h$.

The lower bound uses

$$\sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \geq \frac{1}{2\mu} \|\text{as } \boldsymbol{\tau}\|_{0,\Omega}^2, \tag{3.9}$$

which follows along the same lines as (2.15). The identity (2.15) is replaced by

$$\sum_{T \in \mathcal{T}_h} (\boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T} = -(\text{div } \boldsymbol{\tau}, \mathbf{v})_{0,\Omega} - \sum_{T \in \mathcal{T}_h} (\text{as } \boldsymbol{\tau}, \nabla \mathbf{v})_{0,T}. \tag{3.10}$$

This identity can be shown in the following way:

$$\sum_{T \in \mathcal{T}_h} (\boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T} = \sum_{T \in \mathcal{T}_h} ((\text{sy } \boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T} + (\text{as } \boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T}) = \sum_{T \in \mathcal{T}_h} (\text{sy } \boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T} = \sum_{T \in \mathcal{T}_h} (\text{sy } \boldsymbol{\tau}, \nabla \mathbf{v})_{0,T}$$

$$= \sum_{T \in \mathcal{T}_h} ((\boldsymbol{\tau}, \nabla \mathbf{v})_{0,T} - (\text{as } \boldsymbol{\tau}, \nabla \mathbf{v})_{0,T}) = \sum_{T \in \mathcal{T}_h} ((\boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v})_{0,\partial T} - (\text{div } \boldsymbol{\tau}, \mathbf{v})_{0,T} - (\text{as } \boldsymbol{\tau}, \nabla \mathbf{v})_{0,T}),$$

where

$$\sum_{T \in \mathcal{T}_h} (\boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v})_{0,\partial T} = \sum_{E \subseteq \Gamma_N} (\boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v})_{0,E} + \sum_{E \subseteq \Gamma_D} (\boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v})_{0,E} + \sum_{E \not\subseteq \partial\Omega} (\boldsymbol{\tau} \cdot \mathbf{n}, [\mathbf{v}])_{0,E}$$

and the sum runs over all edges of the triangulation $\mathcal{T}_h$ and $[\mathbf{v}]$ denotes the jump of $\mathbf{v}$ on E. In the above sum the first term vanishes since $\boldsymbol{\tau} \cdot \mathbf{n} = \mathbf{0}$ on Γ_N. For the remaining two terms, we see that $\boldsymbol{\tau} \cdot \mathbf{n}$ is a polynomial of degree at most $k - 1$ and $\mathbf{v}$ or $[\mathbf{v}]$, respectively, is a polynomial of degree k which vanishes at the Gauss points. In both cases, the integrand is therefore a polynomial of degree $2k - 1$ which is zero at the k Gauss points implying that the second and third term also vanish.

Along the same lines as in the proof of Theorem 2.1 we can deduce from (3.9) and (3.10) that

$$\mathcal{B}_h(\boldsymbol{\tau}, \mathbf{v}; \boldsymbol{\tau}, \mathbf{v}) = \|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu \sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{-1/2}\boldsymbol{\tau} - \mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \geq \frac{1}{3}\left(\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2\right.$$

$$+ \mu \sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 + \|\text{as } \boldsymbol{\tau}\|_{0,\Omega}^2 - 2\mu \sum_{T \in \mathcal{T}_h} (\boldsymbol{\tau}, \boldsymbol{\varepsilon}(\mathbf{v}))_{0,T}\left.\right) \geq \frac{1}{6}\|\text{div } \boldsymbol{\tau}\|_{0,\Omega}^2 + \frac{\mu}{6C_K'}\|\mathcal{C}^{-1/2}\boldsymbol{\tau}\|_{0,\Omega}^2$$

$$+ \frac{\mu}{12C_K'} \sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{1/2}\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,T}^2 \geq \frac{1}{12C_K'} \, \||\, (\boldsymbol{\tau}, \mathbf{v}) \,|\|_h^2,$$

where the discrete Korn's inequality (3.6) needs to be used at the appropriate places.   ∎
As a consequence of Theorem 3.1 we have the following quasi-optimality result.

Theorem 3.2. *Let* $\mathbf{X}_h$ *and* $\mathbf{V}_h$ *be as in Theorem* 3.1 *and let* $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h \times \mathbf{V}_h$ *be the solution of* (3.2). *Then there exists a constant C, which is independent of h,* μ *and* λ*, such that*

$$\left|\!\left|\!\left| \, (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h) \, \right|\!\right|\!\right|_h \le c \inf_{(\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h \times \mathbf{V}_h} \left|\!\left|\!\left| \, (\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \mathbf{u} - \mathbf{v}_h) \, \right|\!\right|\!\right|_h. \tag{3.11}$$

Proof. The second lemma of Strang (see, e.g., [12, Section III.1] or [13, Section 10.1]) implies

$$\left|\!\left|\!\left| \, (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h) \, \right|\!\right|\!\right|_h \le c \Bigg(\inf_{(\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h \times \mathbf{V}_h} \left|\!\left|\!\left| \, (\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \mathbf{u} - \mathbf{v}_h) \, \right|\!\right|\!\right|_h$$
$$+ \sup_{(\boldsymbol{\xi}_h, \mathbf{w}_h) \in \mathbf{X}_h \times \mathbf{V}_h} \frac{\mathscr{B}_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h; \boldsymbol{\xi}_h, \mathbf{w}_h)}{\left|\!\left|\!\left| \, (\boldsymbol{\xi}_h, \mathbf{w}_h) \, \right|\!\right|\!\right|_h} \Bigg).$$

The second term in this expression vanishes since

$$\mathscr{B}_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h; \boldsymbol{\xi}_h, \mathbf{w}_h) = (\text{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \text{div } \boldsymbol{\xi}_h)_{0,\Omega} + \mu \sum_{T \in \mathscr{T}_h} (\mathscr{C}^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$$
$$- \mathscr{C}^{1/2}\varepsilon(\mathbf{u} - \mathbf{u}_h), \mathscr{C}^{-1/2}\boldsymbol{\xi}_h - \mathscr{C}^{1/2}\varepsilon(\mathbf{w}_h))_{0,T} = -(\text{div } \boldsymbol{\sigma}_h, \text{div } \boldsymbol{\xi}_h)_{0,\Omega}$$
$$- \mu \sum_{T \in \mathscr{T}_h} (\mathscr{C}^{-1/2}\boldsymbol{\sigma}_h - \mathscr{C}^{1/2}\varepsilon(\mathbf{u}_h), \mathscr{C}^{-1/2}\boldsymbol{\xi}_h - \mathscr{C}^{1/2}\varepsilon(\mathbf{w}_h))_{0,T},$$

which is zero due to (3.2). ∎

For the displacement component of the norm (2.9), we have

$$\inf_{\mathbf{v}_h \in \mathbf{V}_h} \left|\!\left|\!\left| \, (0, \mathbf{u} - \mathbf{v}_h) \, \right|\!\right|\!\right|_h = \inf_{\mathbf{v}_h \in \mathbf{V}_h} \mu \Bigg(\sum_{T \in \mathscr{T}_h} \| \mathscr{C}^{1/2}(\varepsilon(\mathbf{u}) - \varepsilon(\mathbf{v}_h)) \|_{0,T}^2 \Bigg)^{1/2}$$
$$\le \inf_{\mathbf{v}_h \in \mathbf{V}_h} \Bigg(\sum_{T \in \mathscr{T}_h} 2\mu^2 \|\varepsilon(\mathbf{u}) - \varepsilon(\mathbf{v}_h)\|_{0,T}^2 + \lambda\mu \|\text{div}(\mathbf{u} - \mathbf{v}_h)\|_{0,T}^2 \Bigg)^{1/2}. \tag{3.12}$$

Approximation properties with respect to this norm which are uniform in λ can be found in [13, Section 11.4] for the linear case of Crouzeix-Raviart elements. For quadratic nonconforming finite elements such estimates follow along the same lines using results from the original paper of Fortin and Soulie [10] (cf. [15]).

For the stress component of the norm (2.9), we have

$$\inf_{\boldsymbol{\tau}_h \in \mathbf{X}_h} \left|\!\left|\!\left| \, (\boldsymbol{\sigma} - \boldsymbol{\tau}_h, 0) \, \right|\!\right|\!\right|_h = \inf_{\boldsymbol{\tau}_h \in \mathbf{X}_h} (\|\text{div}(\boldsymbol{\sigma} - \boldsymbol{\tau}_h)\|_{0,\Omega}^2 + \mu \|\mathscr{C}^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\tau}_h)\|_{0,\Omega}^2)^{1/2}$$
$$\le \inf_{\boldsymbol{\tau}_h \in \mathbf{X}_h} (\|\text{div}(\boldsymbol{\sigma} - \boldsymbol{\tau}_h)\|_{0,\Omega}^2 + \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{0,\Omega}^2)^{1/2}, \tag{3.13}$$

which is the standard $H(\text{div}, \Omega)$ norm. The approximation properties of Raviart-Thomas elements with respect to this norm can be deduced from [14, Prop. 3.6] and [14, Prop. 3.8].

IV. THE LEAST-SQUARES FUNCTIONAL AS AN *A POSTERIORI* ERROR ESTIMATOR

One of the main motivations for using least-squares finite element approaches is the fact that the element-wise evaluation of the functional serves as an *a posteriori* error estimator. This follows from

$$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) = \|\mathrm{div}\ \boldsymbol{\sigma}_h\|_{0,\Omega}^2 + \mu \sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{-1/2}\boldsymbol{\sigma}_h - \mathcal{C}^{1/2}\varepsilon(\mathbf{u}_h)\|_{0,T}^2 = \|\mathrm{div}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma})\|_{0,\Omega}^2$$

$$+ \mu \sum_{T \in \mathcal{T}_h} \|\mathcal{C}^{-1/2}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}) - \mathcal{C}^{1/2}(\varepsilon(\mathbf{u}_h - \mathbf{u}))\|_{0,T}^2 = \mathcal{B}_h(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \mathbf{u}_h - \mathbf{u}; \boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \mathbf{u}_h - \mathbf{u}),$$

which by Theorem 3.1 is uniformly equivalent to $\vert\!\vert\!\vert (\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \mathbf{u}_h - \mathbf{u}) \vert\!\vert\!\vert_h^2$. For the definition of the *a posteriori* error estimator, we observe that the least-squares functional is the sum of its element-wise contributions,

$$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) = \sum_{T \in \mathcal{T}_h} (\|\mathrm{div}\ \boldsymbol{\sigma}_h\|_{0,T}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\sigma}_h - \mathcal{C}^{1/2}\varepsilon(\mathbf{u}_h)\|_{0,T}^2) =: \sum_{T \in \mathcal{T}_h} \eta_T^2.$$

We collect these observations in the following theorem.

Theorem 4.1. *The element-wise computation of the least-squares functional constitutes an a posteriori error estimator. In other words, for*

$$\eta_T = (\|\mathrm{div}\ \boldsymbol{\sigma}_h\|_{0,T}^2 + \mu\|\mathcal{C}^{-1/2}\boldsymbol{\sigma}_h - \mathcal{C}^{1/2}\varepsilon(\mathbf{u}_h)\|_{0,T}^2)^{1/2},$$

we have

$$\alpha \vert\!\vert\!\vert (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h) \vert\!\vert\!\vert_h^2 \leq \sum_{T \in \mathcal{T}_h} \eta_T^2 \leq \beta \vert\!\vert\!\vert (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h) \vert\!\vert\!\vert_h^2, \tag{4.1}$$

where α and β are the same constants as in Theorem 3.1.

Adaptive refinement strategies consist in refining those triangles with the largest values of η_T. In order to keep the number of degrees of freedom under control, about 20 percent of all triangles were refined regularly (by dividing each in four congruent subtriangles) in the computational experiments reported in the following section. Additionally, irregularly refined triangles are needed in order to make the triangulation admissible. For a detailed investigation of different refinement techniques in connection with the least-squares finite element method; see also [16].

V. COMPUTATIONAL TESTS

In this section, numerical results for a benchmark problem of linear elasticity taken from [3] are presented. The problem to be considered is given by a quadratic membrane of elastic isotropic material with a circular hole in the centre. The upper and lower edges of the strip are loaded with a uniform tensile stress of 4.5 pointing outward. Because of the symmetry of the domain, it suffices to discretize only a fourth of the total geometry. The computational domain is then given by

$$\Omega = \{\mathbf{x} \in \mathbb{R}^2 : 0 < x_1 < 10, 0 < x_2 < 10, x_1^2 + x_2^2 > 1\}$$

(see Fig. 2). The boundary conditions on the top edge of the computational domain ($x_2 = 10$, $0 < x_1 < 10$) are set to $\boldsymbol{\sigma} \cdot \mathbf{n} = (0, 4.5)$, the boundary conditions on the bottom ($x_2 = 0$, 1
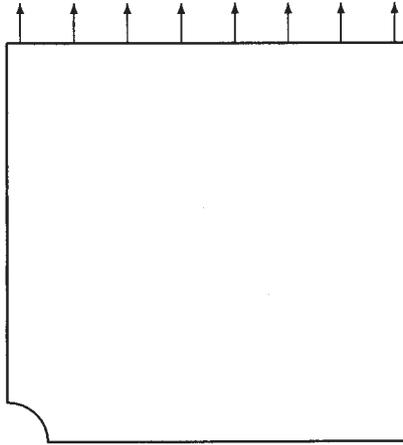
FIG. 2.   Computational domain and boundary conditions.

$< x_1 < 10$) are set to $(\sigma_{11}, \sigma_{12}) \cdot \mathbf{n} = 0$, $u_2 = 0$ (symmetry condition) and finally the boundary conditions on the left ($x_1 = 0$, $1 < x_2 < 10$) are given by $u_1 = 0$, $(\sigma_{21}, \sigma_{22}) \cdot \mathbf{n} = \mathbf{0}$ (symmetry condition). The material parameters are $E = 206900$ for Young's modulus and $\nu = 0.29$ for Poisson's ratio. In addition we also include the results of our numerical tests for $\nu = 0.49$ in order to show the behavior for nearly incompressible materials. Figure 3 shows the triangulation obtained after six adaptive refinement steps.

The following results are computed using the least-squares finite element method on a sequence of adaptively refined meshes based on the *a posteriori* error estimator presented in Section IV. Tables I and II show the results for quadratic nonconforming elements for the displacement approximation combined with quadratic Raviart-Thomas elements for the stress approximation. These tables provide a strong indication that the minimum of the functional is inversely proportional to the square of the number of degrees of freedom,
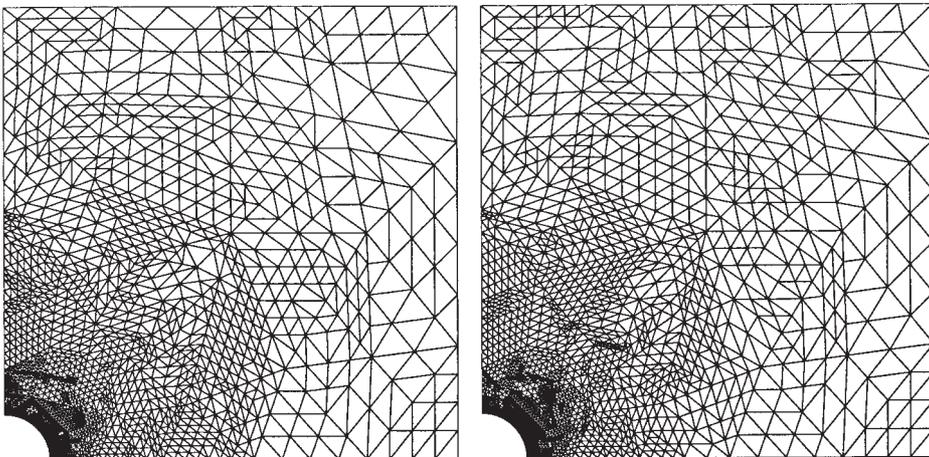

FIG. 3.   Triangulation after six adaptive refinement steps (left: $\nu = 0.29$, right: $\nu = 0.49$).

TABLE I.   Nonconforming quadratic elements for $\mathbf{V}_h$: $\nu = 0.29$.

	# elements	N_X	N_V	$\mathscr{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h)$	$\|\text{as } \boldsymbol{\sigma}_h\|_{0,\Omega}^2$	$\sigma_{22}(1, 0)$
$l = 0$	52	504	328	2.90e-1	2.24e-1	9.946
$l = 1$	117	1150	722	5.02e-2	4.91e-2	12.481
$l = 2$	247	2442	1510	9.68e-3	9.01e-3	13.408
$l = 3$	513	5080	3128	1.91e-3	1.65e-3	13.716
$l = 4$	1063	10544	6464	5.39e-4	4.47e-4	13.824
$l = 5$	2106	20898	12798	1.41e-4	1.12e-4	13.859
$l = 6$	4277	42466	25966	3.95e-5	3.15e-5	13.876
$l = 7$	8465	84070	51370	1.05e-5	8.37e-6	13.881

$$\mathscr{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) \sim \frac{1}{(N_X + N_V)^2}.$$

Also, the square of the $L^2(\Omega)$ norm of the antisymmetric stress part converges at roughly the same rate as the least-squares functional in accordance with the inequality (2.15).

The results in Tables III and IV are obtained with quadratic conforming finite element spaces for the displacement approximation. These tables show the same order of convergence measured by the functional minimum as for the nonconforming case. However, the difference between the nonconforming and conforming quadratic approximation is much larger for $\nu = 0.49$ compared to $\nu = 0.29$. This difference can be clearly seen in the doubly logarithmic convergence graphs in Fig. 4. The gap between the nonconforming (solid line) and conforming (dashed line) approximation is much wider in the graph on the right ($\nu = 0.49$) than it is on the left ($\nu = 0.29$). Using conforming quadratic finite elements for $\mathbf{V}_h$ will therefore not suffice to achieve convergence rates which are uniform in the incompressible limit ($\nu \to \frac{1}{2}$).

Of particular interest in this example is the stress component σ_{22} at the point $(1, 0)$. The size of this stress component is responsible for failure of the material at this point. For $\nu = 0.29$ the value of $\sigma_{22}(1, 0) = 13.8873$ is given in [3] for a reference solution computed by a polynomial approximation of high degree. The corresponding columns in Tables I and III show the convergence of the solutions obtained with our least-squares approach to that reference value as the mesh is refined.

We conclude this section with some remarks comparing our least-squares finite element method to other approaches for linear elasticity. As we have proved in the previous sections and illustrated by the numerical results in Tables I and II, the least-squares mixed finite element method using quadratic Raviart-Thomas elements for the stress and quadratic nonconforming elements for the displacement converges at a rate

TABLE II.   Nonconforming quadratic elements for $\mathbf{V}_h$: $\nu = 0.49$.

	# elements	N_X	N_V	$\mathscr{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h)$	$\|\text{as } \boldsymbol{\sigma}_h\|_{0,\Omega}^2$
$l = 0$	52	504	328	2.40e-1	1.83e-1
$l = 1$	117	1150	722	4.26e-2	4.21e-2
$l = 2$	245	2420	1500	8.30e-3	7.77e-3
$l = 3$	514	5088	3136	1.58e-3	1.41e-3
$l = 4$	1080	10708	6572	4.43e-4	3.64e-4
$l = 5$	2168	21514	13174	1.07e-4	9.16e-5
$l = 6$	4393	43618	26670	3.20e-5	2.61e-5
$l = 7$	8623	85646	52322	8.35e-6	7.21e-6

TABLE III.   Conforming quadratic elements for $\mathbf{V}_h$: $\nu = 0.29$.

	# elements	N_X	N_V	$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h)$	$\|\text{as } \boldsymbol{\sigma}_h\|_{0,\Omega}^2$	$\sigma_{22}(1, 0)$
$l = 0$	52	504	224	5.83e-1	3.61e-1	9.543
$l = 1$	124	1218	518	8.77e-2	6.17e-2	12.508
$l = 2$	262	2590	1078	1.41e-2	1.02e-2	13.435
$l = 3$	565	5600	2310	3.25e-3	2.05e-3	13.737
$l = 4$	1138	11290	4642	7.97e-4	5.14e-4	13.832
$l = 5$	2347	23306	9552	2.24e-4	1.27e-4	13.867
$l = 6$	4643	46124	18878	6.31e-5	3.58e-5	13.878
$l = 7$	9191	91330	37344	1.81e-5	9.49e-6	13.893

$$\left\| \! \left\| \! \left\| \, (\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \mathbf{u}_h - \mathbf{u}) \, \right\| \! \right\| \! \right\|_h \approx \mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h)^{1/2} \sim \frac{1}{N_X + N_V}.$$

This means that we have an optimal order of approximation with respect to an $H(\text{div}, \Omega)^2$ norm for $\boldsymbol{\sigma}$ and an element-wise $H^1(\Omega)^2$ norm for $\mathbf{u}$ (both suitably scaled to account for nearly incompressible materials). The numbers of unknowns involved in this approach is roughly $6N_E + 6N_T + 2N_P$, where N_E, N_T, and N_P are the number of edges, triangles and vertices in the triangulation.

In contrast to this, a method based on a displacement variational formulation using, for example, nonconforming quadratic elements would only involve roughly $2N_E + 2N_T + 2N_P$ unknowns and lead to a comparable approximation for $\mathbf{u}$ which is also uniform in the incompressible limit. Therefore, this method would be preferable in cases where the displacement field is the variable of primary interest. Approximations for the stresses could be computed from the displacement approximations but the convergence would generally be of lower order and only with respect to the $L^2(\Omega)$ norm. In applications where one is mainly interested in an accurate approximation of the stress tensor one should rather rely on a variational formulation involving the stress directly.

Another alternative is to use mixed finite element methods based on the Hellinger-Reissner principle. These methods are constructed with a focus on an accurate stress approximation and treat the displacements as Lagrange parameters. As described in [7], quadratic approximation of the stress could be achieved using the Brezzi-Douglas-Marini elements [17] enriched by bubble functions. In addition, the displacement and an additional set of Lagrange parameters representing the asymmetry of the stress tensor are approximated by piecewise linear and piecewise quadratic finite elements, respectively. The number of degrees of freedoms is comparable to our

TABLE IV.   Conforming quadratic elements for $\mathbf{V}_h$: $\nu = 0.49$.

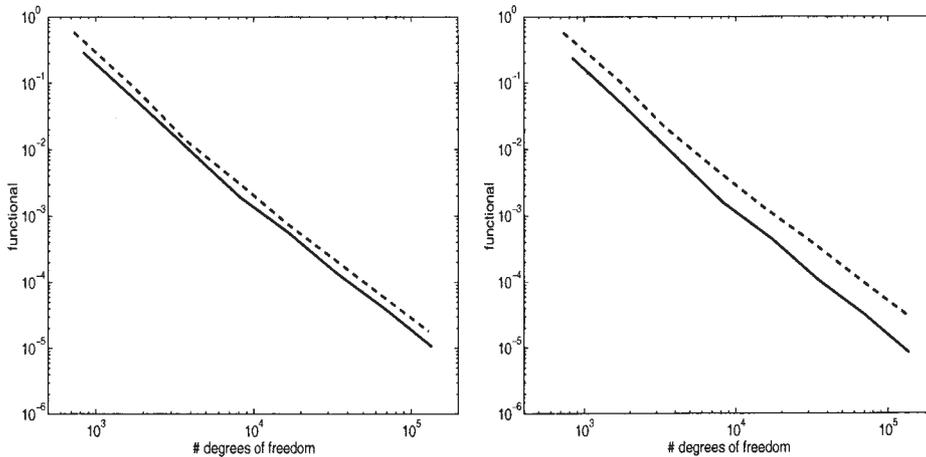|	# elements	N_X	N_V	$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h)$	$\|\text{as } \boldsymbol{\sigma}_h\|_{0,\Omega}^2$
$l = 0$	52	504	224	5.77e-1	3.83e-1
$l = 1$	121	1188	506	1.07e-1	7.06e-2
$l = 2$	244	2410	1006	2.15e-2	1.63e-2
$l = 3$	500	4948	2052	5.49e-3	4.33e-3
$l = 4$	1066	10574	4350	1.38e-3	1.14e-3
$l = 5$	2271	22542	9252	3.90e-4	3.06e-4
$l = 6$	4639	46090	18856	1.10e-4	8.72e-5
$l = 7$	9496	94370	38574	3.14e-5	2.28e-5

FIG. 4.   Nonconforming vs. conforming finite elements for $\mathbf{V}_h$ (left: $\nu = 0.29$, right: $\nu = 0.49$).

least-squares method while the approximation space for the stress is somewhat larger at the cost of a weaker displacement approximation.

These remarks show that the most efficient method for a particular problem is generally dependent on the variables to be approximated and on the norms used to measure the error. An overall comparison would have to incorporate adaptive refinement strategies based on *a posteriori* error estimators for all competing methods. Such error estimators are also available for the mixed methods of Hellinger-Reissner type (see [8] and [9]). However, these *a posteriori* error estimators are certainly more tedious and costly to implement compared to the least-squares approach where the error estimator simply consists of the local evaluation of the functional which is readily available during the computations.

References

 1. Z. Cai and G. Starke, First-order system least squares for the stress-displacement formulation: Linear elasticity, SIAM J Numer Anal 41 (2003), 715–730.

 2. E. Stein, P. Wriggers, A. Rieger, and M. Schmidt, Benchmarks, in Error-controlled Adaptive Finite Elements in Solid Mechanics, E. Stein, Ed., John Wiley and Sons, 2002, Ch. 11, pp. 385–404.

 3. P. B. Bochev and M. D. Gunzburger, Finite element methods of least-squares type, SIAM Rev 40 (1998), 789–837.

 4. B. Jiang, The least-squares finite element method, Springer: Berlin, 1998.

 5. Z. Cai, T. A. Manteuffel, S. F. McCormick, and S. V. Parter, First-order system least squares for planar linear elasticity: Pure traction problem, SIAM J Numer Anal 35 (1998), 320–335.

 6. D. N. Arnold, F. Brezzi, and J. Douglas, PEERS: A new mixed finite element for plane elasticity, Japan J Appl Math 1 (1984), 347–367.

 7. R. Stenberg, A family of mixed finite elements for the elasticity problem, Numer Math 53 (1988), 513–538.

 8. D. Braess, O. Klaas, R. Niekamp, E. Stein, and F. Wobschal, Error indicators for mixed finite elements in 2-dimensional linear elasticity, Comput Methods Appl Mech Engrg 127 (1995), 345–356.

 9. M. Lonsing and R. Verfürth, A posteriori error estimators for mixed finite element methods in linear elasticity, Numer Math, to appear in print, DOI: 10.1007/s00211-004-0519-8.

10. M. Fortin and M. Soulie, A non-conforming piecewise quadratic finite element on triangles, Int J Numer Meth Engrg 19 (1983), 505–520.

11. Z. Cai and G. Starke, Least squares methods for linear elasticity, SIAM J Numer Anal, to appear.

12. D. Braess, Finite elements: Theory, fast solvers, and applications in solid mechanics, Cambridge University Press: Cambridge, 2nd ed., 2001.

13. S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer: New York, 2nd Ed., 2002.

14. F. Brezzi and M. Fortin, Mixed and hybrid finite element methods, Springer: New York, 1991.

15. R. S. Falk, Nonconforming finite element methods for the equations of linear elasticity, Math Comp 57 (1991), 529–550.

16. M. Berndt, T. A. Manteuffel, and S. F. McCormick, Local error estimates and adaptive refinement for first-order system least squares, Electr Trans Numer Anal 6 (1997), 35–43.

17. F. Brezzi, J. Douglas, and L. D. Marini, Two families of mixed finite elements for second order elliptic problems, Numer Math 47 (1985), 217–235.