DE GRUYTER Comput. Methods Appl. Math. 2019; 19(3): 415–430

Research Article

Fleurianne Bertrand*, Zhiqiang Cai and Eun Young Park

Least-Squares Methods for Elasticity and Stokes Equations with Weakly Imposed Symmetry

https://doi.org/10.1515/cmam-2018-0255
Received October 5, 2018; revised February 17, 2019; accepted March 21, 2019

Abstract: This paper develops and analyzes two least-squares methods for the numerical solution of linear elasticity and Stokes equations in both two and three dimensions. Both approaches use the L^2 norm to define least-squares functionals. One is based on the stress-displacement/velocity-rotation/vorticity-pressure (SDRP/SVVP) formulation, and the other is based on the stress-displacement/velocity-rotation/vorticity (SDR/SVV) formulation. The introduction of the rotation/vorticity variable enables us to weakly enforce the symmetry of the stress. It is shown that the homogeneous least-squares functionals are elliptic and continuous in the norm of $H(\mathrm{div};\Omega)$ for the stress, of $H^1(\Omega)$ for the displacement/velocity, and of $L^2(\Omega)$ for the rotation/vorticity and the pressure. This immediately implies optimal error estimates in the energy norm for conforming finite element approximations. As well, it admits optimal multigrid solution methods if Raviart–Thomas finite element spaces are used to approximate the stress tensor. Through a refined duality argument, an optimal L^2 norm error estimates for the displacement/velocity are also established. Finally, numerical results for a Cook's membrane problem of planar elasticity are included in order to illustrate the robustness of our method in the incompressible limit.

Keywords: Least-squares, Linear Elasticity, Stokes Equations

MSC 2010: 65N30

1 Introduction

Least-squares finite element method has been successfully used in computational fluid dynamics, solid mechanics, and electro-magnetics (see, e.g., books by Bochev and Gunzburger [10] and by Jiang [24], and references therein). The method has many attractions. The two striking features are (i) it naturally symmetrizes and stabilizes the original problem; and (ii) the corresponding least-squares functional is an accurate a posteriori error estimator. Stability plays a crucial role for numerical algorithms, and symmetry is important for developing fast solvers. Accurate error estimator guarantees reliability of the computation and efficiency of adaptive mesh refinement algorithm.

For linear elasticity problems, least-squares finite element methods have been studied based on various first-order system, e.g., the displacement-displacement gradient formulation in [18], the stress-displacement-rotation formulation in [25] (see also the references therein for some other least-squares approaches in the engineering literature), and the stress-displacement formulation in [19, 20]. One drawback of the approaches

***Corresponding author: Fleurianne Bertrand,** Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, e-mail: fb@math.hu-berlin.de
Zhiqiang Cai, Eun Young Park, Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA, e-mail: caiz@purdue.edu, park296@purdue.com

in [18, 25] is their requirement of sufficient smoothness on the original problem if using standard continuous finite element approximations. This drawback was overcome in [19, 20] by applying the L^2 norm least squares to the stress-displacement formulation. Moreover, it was shown in [20] that the homogeneous least-squares functional is elliptic and continuous in the $H(\mathrm{div};\Omega)^d \times H^1(\Omega)^d$ norm, $d = 2$ or 3, uniform in the incompressible limit. This implies optimal error estimates in the energy norm for conforming finite element approximations as well as optimal multigrid solution methods if Raviart–Thomas finite element spaces are used to approximate the stress tensor.

Recently, a "least-squares" finite element method was introduced in [27] from an engineering perspective and analyzed in [28]. This method modifies the least squares approach in [20] by adding the skew-symmetric part of displacement gradient in the test space. It was shown numerically that the modified method improves momentum balance and gives better results in bending dominated situations. However, the resulting variational formulation is non-symmetric; the method is no longer of the least-squares type and, hence, it does not have a natural error estimator.

Motivated by works in [27, 28], in this paper we study the L^2 norm least-squares finite element method based on the stress-displacement-rotation (SDR) formulation (see (2.7)) with a hope of improving momentum balance. The rotation is defined as the curl of the displacement field, and the SDR formulation in this paper, that is different from that in [25], has been used to develop mixed finite element methods (see, e.g., [1, 3, 23] and [4]). The degree of freedoms for mixed elements based on the SDR formulation is much less than that of mixed elements based on the stress-displacement formulation (see, e.g., [2, 5]). This is because the symmetry condition of the stress is imposed weakly in the SDR formulation. By employing the least-squares principle, the number of the degrees of freedom is further reduced. For the stress-displacement formulation, see [20]. For the stress-displacement-rotation (SDR) formulation in two dimensions, for example, the lowest order finite element spaces for the least-squares method introduced in this paper and for the PEERs in [3] are the respective $\mathrm{RT}_0^2 \times P_1^2 \times P_0$ and $(\mathrm{RT}_0^2 + B) \times P_0^2 \times P_1$, where B is the span of the gradient perp of the cubic bubble functions. Hence, their degrees of freedom are 5 and 7.5 per element, where edge and vertex freedoms are counted as half and one-sixth, respectively.

To analyze the least-squares method, the key step is to establish the coercivity of the homogeneous least-squares functional. For the least-squares method based on the SDR formulation, we are not able to directly prove the coercivity bound. This is because the constitutive equation involves all three variables, but the other two equations (the equilibrium equation and the symmetry constraint) have only the stress variables, To circumvent this difficulty, we introduce a new variable, the hydrostatic pressure, defined as the average of the normal stresses. Instead of using the definition of the pressure as the new equation, we derive an equation connecting the pressure and the displacement through the trace of the constitutive equation. The resulting first-order system is called the stress-displacement-rotation-pressure (SDRP) formulation (see (2.9)).

The main theoretical result of the paper is to establish the coercivity of the homogeneous least-squares functional based on the SDRP formulation (see Theorem 1) uniform with respect to the Lamé constant in the norm of $H(\mathrm{div};\Omega)$ for the stress, of $H^1(\Omega)$ for the displacement, and of $L^2(\Omega)$ for the rotation and the pressure. As a direct consequence, the coercivity of the homogeneous least-squares functional based on the SDR formulation follows easily. With the coercivity and the continuity, it is then easy to obtain optimal error estimates of the least-squares finite element methods using conforming finite element approximations in the energy norm. As well, the resulting algebraic system may be solved numerically by optimal multigrid solution methods if Raviart–Thomas finite element spaces are used to approximate the stress tensor. Moreover, through a refined duality argument introduced in [14], we are able to obtain an optimal L^2 norm error estimates for the displacement. Finally, numerical results for Cook's membrane problem of planar elasticity are included in order to illustrate the robustness of our method in the incompressible limit.

For incompressible Stokes equation, least-squares finite element methods based on various formulations and various norms have been proposed, analyzed, implemented, and tested (see, e.g., [7–9, 12, 16, 17, 21, 24]). In particular, the least-squares methods based on the stress-velocity and the stress-velocity-pressure formulations were studied in [15]. The stress-velocity formulation is identical to the stress-displacement formulation of linear elasticity corresponding to the incompressible limit, i.e., the Lamé constant $\lambda = +\infty$. This means that least-squares methods studied in this paper may be directly applied to the Stokes equation with different physical quantities, where the displacement and the rotation of solids are replaced by the velocity and the vorticity of fluids, respectively.

This least-squares method is closely related to the Hellinger–Reissner mixed formulation, see [11]. The attractions mentioned at the beginning of this introduction remains valid. The approximations of the constrained, i.e., the momentum balance and the symmetry of the stress tensor are super-closed.

This paper is organized as follows. The SDR (SVV) and SDRP (SVVP) formulations for linear elasticity (Stokes) equations are introduced in Section 2. Least-squares minimization problems based on both the SDR and SDRP formulations are analyzed in Section 3 by establishing the coercivity and continuity of the corresponding homogeneous least-squares functionals. In Section 4, we derive optimal error bounds of least-squares finite element approximations in the energy norm as well as the L^2 norm for the displacement/velocity. Finally, numerical experiments for Cook's membrane problem of linear elasticity are presented in Section 5.

1.1 Notations

Denote by $\boldsymbol{\delta} = \boldsymbol{\delta}_{d\times d}$ the identity matrix. Let

$$\boldsymbol{\chi} = \begin{cases} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \text{if } d = 2, \\ (\boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \boldsymbol{\chi}_3) & \text{if } d = 3, \end{cases}$$

where $\boldsymbol{\chi}_i$ for $i = 1, 2, 3$ are 3×3 matrices given by

$$\boldsymbol{\chi}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{\chi}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\chi}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let $\boldsymbol{\sigma} = (\sigma_{ij})_{d\times d}$ and $\boldsymbol{\tau} = (\tau_{ij})_{d\times d}$ be matrix-valued functions in $\mathbb{R}^{d\times d}$, and denote the component-wise dot product by

$$\boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{i,j=1}^{d} \sigma_{ij}\tau_{ij}.$$

Note that if $\boldsymbol{\sigma}$ is symmetric and $\boldsymbol{\tau}$ is skew-symmetric, then

$$\boldsymbol{\sigma} : \boldsymbol{\tau} = 0.$$

The divergence of a tensor $\boldsymbol{\tau}$ is defined by applying the divergence operator to rows by

$$\nabla \cdot \boldsymbol{\tau} = \left(\sum_{i=1}^{d} \frac{\partial \tau_{1,i}}{\partial x_i}, \dots, \sum_{i=1}^{d} \frac{\partial \tau_{d,i}}{\partial x_i} \right)^t.$$

The trace of a tensor $\boldsymbol{\tau}$ is defined by

$$\operatorname{tr} \boldsymbol{\tau} = \tau_{11} + \cdots + \tau_{dd}.$$

Let $\mathbf{v} = (v_1, \dots, v_d)^t$ be a vector-valued function in $\mathbb{R}^d$; the gradient and curl of $\mathbf{v}$ are given by

$$\nabla \mathbf{v} = \begin{pmatrix} \partial_1 v_1 & \cdots & \partial_d v_1 \\ \vdots & \ddots & \vdots \\ \partial_1 v_d & \cdots & \partial_d v_d \end{pmatrix} \quad \text{and} \quad \nabla \times \mathbf{v} = \begin{cases} \nabla \mathbf{v} : \boldsymbol{\chi} & \text{if } d = 2, \\ (\nabla \mathbf{v} : \boldsymbol{\chi}_1, \nabla \mathbf{v} : \boldsymbol{\chi}_2, \nabla \mathbf{v} : \boldsymbol{\chi}_3)^t & \text{if } d = 3. \end{cases}$$

Define the dot product between a vector $\mathbf{v}$ and a tensor $\boldsymbol{\tau}$ by

$$\mathbf{v} \cdot \boldsymbol{\tau} = \left(\sum_{i=1}^{d} v_i \tau_{1,i}, \dots, \sum_{i=1}^{d} v_i \tau_{d,i} \right)^t.$$

2 Elasticity and Stokes Equations

In this section, we first describe the stress-displacement (velocity) formulation for elasticity (Stokes) equations. By introducing independent variables: vorticity and pressure, we then derive the stress-displacement/velocity-vorticity and the stress-displacement/velocity-vorticity-pressure formulations.

Let Ω be a bounded, open, connected domain in $\mathbb{R}^d$ $(d = 2, 3)$ with a Lipschitz continuous boundary $\partial\Omega$. Assume that $\partial\Omega$ consists of two open subsets Γ_D and Γ_N such that $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial\Omega$ and $\Gamma_D \cap \Gamma_N \neq \emptyset$. For simplicity, we assume $\Gamma_D \neq \emptyset$. Let $\mathbf{u} = (u_1, \dots, u_d)^t$ be the displacement and the velocity for the elasticity and Stokes equations, respectively. Given external force $\mathbf{f} = (f_1, \dots, f_d)^t$, denote the stress tensor by $\boldsymbol{\sigma} = (\sigma_{ij})_{d\times d}$; then we have the following equilibrium equation:

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{f} \quad \text{in } \Omega.$$

Denote by $\boldsymbol{\epsilon}(\mathbf{u}) = (\epsilon_{ij}(\mathbf{u}))_{d\times d}$ the strain tensor; then

$$\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^t). \tag{2.1}$$

The relation between the stress and the strain tensors is represented as follows by the constitutive law:

$$\mathcal{A}_\lambda \boldsymbol{\sigma} = \epsilon(\mathbf{u}),$$

where $\mathcal{A}_\lambda$ is a fourth-order tensor. For an isotropic elastic material, $\mathcal{A}_\lambda$ is called the compliance tensor given by

$$\mathcal{A}_\lambda \boldsymbol{\tau} = \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{\lambda}{2\mu + d\lambda}(\operatorname{tr}\boldsymbol{\tau})\boldsymbol{\delta}\right) \tag{2.2}$$

for any tensor $\boldsymbol{\tau} = (\tau_{ij})_{d\times d}$, where μ and λ are the Lamé constants such that $\mu \in [\mu_1, \mu_2]$ with $0 < \mu_1 < \mu_2 < \infty$ and $0 < \lambda \leq \infty$. The material is said to be nearly incompressible or incompressible when λ is very large or infinite, respectively. As the λ approaches ∞, the compliance tensor tends to

$$\mathcal{A}\boldsymbol{\tau} = \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{1}{d}(\operatorname{tr}\boldsymbol{\tau})\boldsymbol{\delta}\right), \tag{2.3}$$

which is not invertible. For incompressible Newtonian fluids, (2.3) also holds, where the μ is the viscosity constant. Hence, the first-order system for the stress and the displacement/velocity is expressed as follows:

$$\begin{cases} \mathcal{A}_\lambda\boldsymbol{\sigma} - \boldsymbol{\epsilon}(\mathbf{u}) = \mathbf{0} & \text{in } \Omega, \\ -\nabla\cdot\boldsymbol{\sigma} = \mathbf{f} & \text{in } \Omega. \end{cases}$$

This system is closed with the following (for simplicity) homogeneous boundary conditions:

$$\begin{cases} \mathbf{u} = \mathbf{0} & \text{on } \Gamma_D, \\ \mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{0} & \text{on } \Gamma_N, \end{cases} \tag{2.4}$$

where $\mathbf{n} = (n_1, \dots, n_d)^t$ is the outward unit vector normal to the boundary $\partial\Omega$.

By (2.1), the strain tensor is the symmetric part of $\nabla\mathbf{u}$. Hence, it may be rewritten as the difference of $\nabla\mathbf{u}$ and its skew-symmetric part, i.e.,

$$\boldsymbol{\epsilon}(\mathbf{u}) = \nabla\mathbf{u} - \frac{1}{2}(\nabla\mathbf{u} - \nabla\mathbf{u}^t) = \nabla\mathbf{u} - (-1)^d\boldsymbol{\omega}\cdot\boldsymbol{\chi}, \tag{2.5}$$

where $\boldsymbol{\omega} = \frac{1}{2}\nabla\times\mathbf{u}$ denotes the vorticity, which is a scalar for $d = 2$ and a vector for $d = 3$, and $\boldsymbol{\omega}\cdot\boldsymbol{\chi}$ is defined by

$$\boldsymbol{\omega}\cdot\boldsymbol{\chi} = \begin{cases} \omega\boldsymbol{\chi}, & d = 2, \\ (\boldsymbol{\omega}\cdot\boldsymbol{\chi}_1, \boldsymbol{\omega}\cdot\boldsymbol{\chi}_2, \boldsymbol{\omega}\cdot\boldsymbol{\chi}_3), & d = 3. \end{cases}$$

Note that simple calculation gives

$$(\boldsymbol{\omega}\cdot\boldsymbol{\chi}) : \boldsymbol{\tau} = (-1)^d\boldsymbol{\omega}\cdot(\boldsymbol{\chi} : \boldsymbol{\tau}) \quad \text{and} \quad (\boldsymbol{\omega}\cdot\boldsymbol{\chi}) : \boldsymbol{\chi} = (-1)^d 2\boldsymbol{\omega}. \tag{2.6}$$

The constitute equation becomes

$$\mathcal{A}_\lambda \boldsymbol{\sigma} - \nabla \mathbf{u} + (-1)^d \boldsymbol{\omega} \cdot \boldsymbol{\chi} = \mathbf{0} \quad \text{in } \Omega.$$

Since the vorticity is an independent variable, the above equation no longer implies the symmetry of the stress. Hence, we need to impose the following symmetry condition:

$$\operatorname{as} \boldsymbol{\sigma} = \mathbf{0} \quad \text{in } \Omega,$$

where $\operatorname{as} \boldsymbol{\sigma}$ denotes the skew-symmetric part of the $\boldsymbol{\sigma}$ given by

$$\operatorname{as} \boldsymbol{\sigma} = \boldsymbol{\sigma} : \boldsymbol{\chi} = \begin{cases} \sigma_{21} - \sigma_{12}, & d = 2, \\ (\sigma_{32} - \sigma_{23}, \sigma_{13} - \sigma_{31}, \sigma_{21} - \sigma_{12})^t, & d = 3. \end{cases}$$

Now, we have the following stress-displacement/velocity-rotation/vorticity (SDR/SVV) formulation:

$$\begin{cases} \mathcal{A}_\lambda \boldsymbol{\sigma} - \nabla \mathbf{u} + (-1)^d \boldsymbol{\omega} \cdot \boldsymbol{\chi} = \mathbf{0} & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = \mathbf{f} & \text{in } \Omega, \\ \operatorname{as} \boldsymbol{\sigma} = \mathbf{0} & \text{in } \Omega \end{cases} \tag{2.7}$$

with the boundary conditions given in (2.4).

In the remainder of this section, we derive the stress-displacement/velocity-rotation/vorticity-pressure (SDRP/SVVP) formulation. To this end, introducing the hydrostatic pressure

$$p = -\frac{1}{d} \operatorname{tr} \boldsymbol{\sigma}, \tag{2.8}$$

then the constitutive equation becomes

$$\frac{1}{2\mu}\left(\boldsymbol{\sigma} + \frac{d\lambda}{2\mu + d\lambda} p \boldsymbol{\delta}\right) - \nabla \mathbf{u} + (-1)^d \boldsymbol{\omega} \cdot \boldsymbol{\chi} = \mathbf{0} \quad \text{in } \Omega.$$

Applying the trace operator to the above equation and using (2.8), we have

$$\frac{d}{2\mu + d\lambda} p + \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega.$$

Then the SDRP (SVVP) formulation is as follows:

$$\begin{cases} \frac{1}{2\mu} \boldsymbol{\sigma} + \frac{d\lambda}{2\mu(2\mu + d\lambda)} p \boldsymbol{\delta} - \nabla \mathbf{u} + (-1)^d \boldsymbol{\omega} \cdot \boldsymbol{\chi} = \mathbf{0} & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = \mathbf{f} & \text{in } \Omega, \\ \operatorname{as} \boldsymbol{\sigma} = \mathbf{0} & \text{in } \Omega, \\ \frac{d}{2\mu + d\lambda} p + \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega \end{cases} \tag{2.9}$$

with the boundary conditions given in (2.4).

3 Least-Squares Variational Formulation

We use the standard notation and definition for the Sobolev spaces $H^s(\Omega)$ for $s \geq 0$. The inner product and norm are denoted by $(\cdot\,,\cdot)_{s,\Omega}$ and $\|\cdot\|_{s,\Omega}$, respectively. If $s = 0$, then $H^s(\Omega) = L^2(\Omega)$ and we drop the subscript in the notation of the inner product and norm. Let

$$H_D^1(\Omega)^d = \{\mathbf{v} \in H^1(\Omega)^d : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D\}$$

and

$$H(\mathrm{div};\Omega)^d = \{\boldsymbol{\tau} \in L^2(\Omega)^{d\times d} : \nabla\cdot\boldsymbol{\tau} \in L^2(\Omega)^d\},$$

which is a Hilbert space with respect to the following norm:

$$\|\boldsymbol{\tau}\|_{H(\mathrm{div};\Omega)} = (\|\boldsymbol{\tau}\|^2 + \|\nabla\cdot\boldsymbol{\tau}\|^2)^{\frac{1}{2}}.$$

In this section, we introduce least-squares problems for both the SDR (SVV) and the SDRP (SVVP) formulations and show their well-posedness.

Before applying the least-squares principle to the first-order systems introduced in the previous section, we first describe solution spaces. Let

$$\Sigma = \begin{cases} H(\mathrm{div};\Omega)^d & \text{if } \Gamma_N \neq \emptyset, \\ \Big\{\boldsymbol{\tau} \in H(\mathrm{div};\Omega)^d : \int_\Omega \operatorname{tr}\boldsymbol{\tau}\,dx = 0\Big\} & \text{otherwise,} \end{cases}$$

and denote its subspace by

$$\Sigma_N = \{\boldsymbol{\tau} \in \Sigma : \mathbf{n}\cdot\boldsymbol{\tau} = \mathbf{0} \text{ on } \Gamma_N\}.$$

Let

$$\bar{L}^2(\Omega) = \begin{cases} L^2(\Omega) & \text{if } \Gamma_N \neq \emptyset, \\ \Big\{\boldsymbol{\gamma} \in L^2(\Omega) : \int_\Omega \boldsymbol{\gamma}\,dx = 0\Big\} & \text{otherwise,} \end{cases}$$

and let

$$\mathcal{V} = \Sigma_N \times H_D^1(\Omega)^d \times \bar{L}^2(\Omega)^{2d-3} \quad \text{and} \quad \mathcal{W} = \mathcal{V} \times \bar{L}^2(\Omega).$$

For $\mathbf{f} \in L^2(\Omega)^d$ and $\lambda \in (0,\infty]$, we define the following least-squares functionals:

$$G(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma};\mathbf{f}) = \|\mathcal{A}_\lambda\boldsymbol{\tau} - \nabla\mathbf{v} + (-1)^d\boldsymbol{\gamma}\cdot\boldsymbol{\chi}\|^2 + \|\nabla\cdot\boldsymbol{\tau} + \mathbf{f}\|^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|^2 \tag{3.1}$$

for all $(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma}) \in \mathcal{V}$ based on the SDR (SVV) formulation and

$$F(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q;\mathbf{f}) = \left\|\frac{1}{2\mu}\Big(\boldsymbol{\tau} + \frac{d\lambda}{2\mu+d\lambda}q\boldsymbol{\delta}\Big) - \nabla\mathbf{v} + (-1)^d\boldsymbol{\gamma}\cdot\boldsymbol{\chi}\right\|^2 + \|\nabla\cdot\boldsymbol{\tau} + \mathbf{f}\|^2 + \|\mathrm{as}\,\boldsymbol{\tau}\|^2 + \left\|\frac{d}{2\mu+d\lambda}q + \nabla\cdot\mathbf{v}\right\|^2$$

for all $(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q) \in \mathcal{W}$ based on the SDRP (SVVP) formulation. The corresponding least-squares minimization problems are, respectively, defined as follows: finding $(\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\omega}) \in \mathcal{V}$ such that

$$G(\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\omega};\mathbf{f}) = \inf_{(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma})\in\mathcal{V}} G(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma};\mathbf{f}), \tag{3.2}$$

and finding $(\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\omega},p) \in \mathcal{W}$ such that

$$F(\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\omega},p;\mathbf{f}) = \inf_{(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q)\in\mathcal{W}} F(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q;\mathbf{f}). \tag{3.3}$$

For any $(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma}) \in \mathcal{V}$ and any $(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q) \in \mathcal{W}$, define the following energy norms:

$$|||(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma})|||_{\mathcal{V}}^2 \equiv \|\boldsymbol{\tau}\|_{H(\mathrm{div};\Omega)}^2 + \|\mathbf{v}\|_1^2 + \|\boldsymbol{\gamma}\|^2$$

and

$$|||(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q)|||_{\mathcal{W}}^2 \equiv |||(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma})|||^2 + \|q\|^2,$$

respectively. In this paper, we use C with or without subscripts to denote a generic positive constants, possibly different at different occurrences, which is independent of the Lamé constant $\lambda \in (0,\infty]$ and the mesh size h introduced in Section 4 but may depend on the domain Ω. Note that one could scale the variables and the right-hand side accordingly so that μ is equal to one.

Lemma 1. *There exists a positive constant C such that*

$$\left\| \frac{1}{2\mu}\left(\operatorname{tr} \boldsymbol{\tau} + \frac{d^2\lambda}{2\mu + d\lambda} q \right) - \nabla \cdot \mathbf{v} \right\| \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}} \tag{3.4}$$

and

$$\left\| \frac{1}{2\mu} \operatorname{as} \boldsymbol{\tau} - \nabla \times \mathbf{v} + 2\boldsymbol{\gamma} \right\| \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}} \tag{3.5}$$

for all $(\boldsymbol{\tau}, \mathbf{v}, q, \boldsymbol{\gamma}) \in \mathcal{W}$.

Proof. For all $(\boldsymbol{\tau}, \mathbf{v}, q, \boldsymbol{\gamma}) \in \mathcal{W}$, a simple calculation gives

$$\frac{1}{2\mu}\left(\operatorname{tr} \boldsymbol{\tau} + \frac{d^2\lambda}{2\mu + d\lambda} q \right) - \nabla \cdot \mathbf{v} = \left(\frac{1}{2\mu}\left(\boldsymbol{\tau} + \frac{d\lambda}{2\mu + d\lambda} q\boldsymbol{\delta} \right) - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi} \right) : \boldsymbol{\delta}.$$

Now, (3.4) is an immediate consequence of the Cauchy–Schwarz inequality:

$$\left\| \frac{1}{2\mu}\left(\operatorname{tr} \boldsymbol{\tau} + \frac{d^2\lambda}{2\mu + d\lambda} q \right) - \nabla \cdot \mathbf{v} \right\| \le \sqrt{d} \left\| \frac{1}{2\mu}\left(\boldsymbol{\tau} + \frac{d\lambda}{2\mu + d\lambda} q\boldsymbol{\delta} \right) - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi} \right\| \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}}.$$

Estimate (3.5) may be proved in a similar fashion by noting (2.6) and that

$$\frac{1}{2\mu} \operatorname{as} \boldsymbol{\tau} - \nabla \times \mathbf{v} + 2\boldsymbol{\gamma} = \left(\frac{1}{2\mu}\left(\boldsymbol{\tau} + \frac{d\lambda}{2\mu + d\lambda} q\boldsymbol{\delta} \right) - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi} \right) : \boldsymbol{\chi}.$$

This completes the proof of the lemma. □

Lemma 2. *For all $\mathbf{v} \in H^1_D(\Omega)^d$, there exists a positive constant C such that*

$$\|\nabla \mathbf{v}\|^2 \le C\left(\|\nabla \mathbf{v}\|^2 - \frac{1}{2}\|\nabla \times \mathbf{v}\|^2 \right). \tag{3.6}$$

Proof. A simple calculation gives

$$\|\epsilon(\mathbf{v})\|^2 = \|\nabla \mathbf{v}\|^2 - \frac{5-d}{4}\|\nabla \times \mathbf{v}\|^2 \le \|\nabla \mathbf{v}\|^2 - \frac{1}{2}\|\nabla \times \mathbf{v}\|^2.$$

Estimate (3.6) is then a direct consequence of Korn's inequality. □

Lemma 3. *For all $\boldsymbol{\tau} \in \Sigma_N$, there exists a positive constant C such that*

$$\|\boldsymbol{\tau}\|^2 \le C(\|\mathcal{A}\boldsymbol{\tau}\|^2 + \|\nabla \cdot \boldsymbol{\tau}\|^2). \tag{3.7}$$

Proof. It is shown in [19]. □

The next theorem shows that the homogeneous least-squares functionals $G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0})$ and $F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})$ are equivalent to the energy norms.

Theorem 1. *Independent of the Lamé constant $\lambda \in (0, \infty]$, we have the following estimates:*

$$\frac{1}{C}|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q)|||^2_{\mathcal{W}} \le F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0}) \le C|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q)|||^2_{\mathcal{W}} \tag{3.8}$$

for all $(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) \in \mathcal{W}$ and

$$\frac{1}{C}|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma})|||^2_{\mathcal{V}} \le G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0}) \le C|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma})|||^2_{\mathcal{V}} \tag{3.9}$$

for all $(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) \in \mathcal{V}$.

Proof. The upper bound in (3.8) is an immediate consequence of the triangle inequality, and the upper bound in (3.9) follows from the triangle inequality and the fact that

$$\|\mathcal{A}_\lambda \boldsymbol{\tau}\|^2 = \frac{1}{4\mu^2}\left(\|\boldsymbol{\tau}\|^2 - \frac{\lambda(d\lambda + 4\mu)}{(d\lambda + 2\mu)^2}\|\operatorname{tr} \boldsymbol{\tau}\|^2 \right) \le \frac{1}{4\mu^2}\|\boldsymbol{\tau}\|^2.$$

By the definitions of the least-squares functionals and the triangle inequality, we have

$$
\begin{aligned}
F\Big(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, -\frac{1}{d}\operatorname{tr}\boldsymbol{\tau}; \mathbf{0}\Big) &= G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0}) + \Big\|\nabla\cdot\mathbf{v} - \frac{1}{2\mu + d\lambda}\operatorname{tr}\boldsymbol{\tau}\Big\|^2 \\
&= G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0}) + \|\operatorname{tr}(\mathcal{A}_\lambda\boldsymbol{\tau} - \nabla\mathbf{v} + (-1)^d\boldsymbol{\gamma}\cdot\boldsymbol{\chi})\|^2 \le CG(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0}).
\end{aligned}
$$

Hence, the lower bound in (3.9) follows from the lower bound of (3.8) with $q = -\frac{1}{d}\operatorname{tr}\boldsymbol{\tau} \in L^2_D(\Omega)$.

To show the validity of the lower bound in (3.8), using the facts that

$$
\boldsymbol{\chi} : \nabla\mathbf{v} = \nabla\times\mathbf{v} \quad \text{and} \quad \boldsymbol{\delta} : \nabla\mathbf{v} = \nabla\cdot\mathbf{v},
$$

equations (2.6), and integration by parts, we have

$$
\begin{aligned}
\|\nabla\mathbf{v}\|^2 &\le \|\nabla\mathbf{v}\|^2 + \frac{d^2\lambda}{2\mu(2\mu+d\lambda)^2}\|q\|^2 \\
&= \Big(\nabla\mathbf{v} - (-1)^d\boldsymbol{\gamma}\cdot\boldsymbol{\chi} - \frac{1}{2\mu}\Big(\boldsymbol{\tau} + \frac{d\lambda}{2\mu+d\lambda}q\boldsymbol{\delta}\Big), \nabla\mathbf{v}\Big) + (\boldsymbol{\gamma}, \nabla\times\mathbf{v}) - \frac{1}{2\mu}(\nabla\cdot\boldsymbol{\tau}, \mathbf{v}) \\
&\quad + \frac{d\lambda}{2\mu(2\mu+d\lambda)}\Big(q, \nabla\cdot\mathbf{v} + \frac{d}{2\mu+d\lambda}q\Big) \\
&\le \Big\|\nabla\mathbf{v} - (-1)^d\boldsymbol{\gamma}\cdot\boldsymbol{\chi} - \frac{1}{2\mu}\Big(\boldsymbol{\tau} + \frac{d\lambda}{2\mu+d\lambda}q\boldsymbol{\delta}\Big)\Big\|\|\nabla\mathbf{v}\| + |(\boldsymbol{\gamma}, \nabla\times\mathbf{v})| + C\|\nabla\cdot\boldsymbol{\tau}\|\|\mathbf{v}\| \\
&\quad + C\Big\|\nabla\cdot\mathbf{v} + \frac{d}{2\mu+d\lambda}q\Big\|\|q\|,
\end{aligned}
$$

which, together with the Poincaré inequality, implies

$$
\|\nabla\mathbf{v}\|^2 \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}}(\|\nabla\mathbf{v}\| + \|q\|) + |(\boldsymbol{\gamma}, \nabla\times\mathbf{v})|. \tag{3.10}
$$

By the triangle and the Cauchy–Schwarz inequalities and (3.5), the second term on the right-hand side of (3.10) may be bounded above as follows:

$$
\begin{aligned}
|(\boldsymbol{\gamma}, \nabla\times\mathbf{v})| &= \Big|\frac{1}{2}\Big(2\boldsymbol{\gamma} + \frac{1}{2\mu}\operatorname{as}\boldsymbol{\tau} - \nabla\times\mathbf{v}, \nabla\times\mathbf{v}\Big) - \frac{1}{4\mu}(\operatorname{as}\boldsymbol{\tau}, \nabla\times\mathbf{v}) + \frac{1}{2}\|\nabla\times\mathbf{v}\|^2\Big| \\
&\le \Big(\frac{1}{2}\Big\|2\boldsymbol{\gamma} + \frac{1}{2\mu}\operatorname{as}\boldsymbol{\tau} - \nabla\times\mathbf{v}\Big\| + \frac{1}{4\mu}\|\operatorname{as}\boldsymbol{\tau}\|\Big)\|\nabla\times\mathbf{v}\| + \frac{1}{2}\|\nabla\times\mathbf{v}\|^2 \\
&\le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}}\|\nabla\times\mathbf{v}\| + \frac{1}{2}\|\nabla\times\mathbf{v}\|^2,
\end{aligned}
$$

which, together with (3.10), implies

$$
\|\nabla\mathbf{v}\|^2 \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}}(\|\nabla\mathbf{v}\| + \|q\|) + \frac{1}{2}\|\nabla\times\mathbf{v}\|^2.
$$

Combining with Lemma 2, we have

$$
\|\nabla\mathbf{v}\|^2 \le C\Big(\|\nabla\mathbf{v}\|^2 - \frac{1}{2}\|\nabla\times\mathbf{v}\|^2\Big) \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}}(\|\nabla\mathbf{v}\| + \|q\|). \tag{3.11}
$$

To bound the L^2 norm of q, we rewrite q as follows:

$$
q = \frac{1}{d}\Big(\frac{d^2\lambda}{2\mu+d\lambda}q + \operatorname{tr}\boldsymbol{\tau} - 2\mu\nabla\cdot\mathbf{v}\Big) + \frac{2\mu}{d}\Big(\nabla\cdot\mathbf{v} + \frac{d}{2\mu+d\lambda}q\Big) - \frac{1}{d}\operatorname{tr}\boldsymbol{\tau}.
$$

It then follows from the triangle inequality and (3.4) that

$$
\begin{aligned}
\|q\| &\le \frac{1}{d}\Big\|\operatorname{tr}\boldsymbol{\tau} + \frac{d^2\lambda}{2\mu+d\lambda}q - 2\mu\nabla\cdot\mathbf{v}\Big\| + \frac{2\mu}{d}\Big\|\nabla\cdot\mathbf{v} + \frac{d}{2\mu+d\lambda}q\Big\| + \frac{1}{d}\|\operatorname{tr}\boldsymbol{\tau}\| \\
&\le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}} + \|\boldsymbol{\tau}\|).
\end{aligned} \tag{3.12}
$$

Next, we use Lemma 3 to bound the L^2 norm of $\boldsymbol{\tau}$. To this end, by the facts that

$$\boldsymbol{\delta} : \mathcal{A}\boldsymbol{\tau} = \operatorname{tr}(\mathcal{A}\boldsymbol{\tau}) = 0 \quad \text{and} \quad \boldsymbol{\chi} : \mathcal{A}\boldsymbol{\tau} = \frac{1}{2\mu} \operatorname{as} \boldsymbol{\tau},$$

and (2.6), a simple calculation gives

$$\|\mathcal{A}\boldsymbol{\tau}\|^2 = \frac{1}{2\mu}(\boldsymbol{\tau}, \mathcal{A}\boldsymbol{\tau}) = \Big(\frac{1}{2\mu}\Big(\boldsymbol{\tau} + \frac{d\lambda}{(2\mu + d\lambda)} q\boldsymbol{\delta}\Big) - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi}, \mathcal{A}\boldsymbol{\tau}\Big) + (\nabla \mathbf{v}, \mathcal{A}\boldsymbol{\tau}) - \frac{1}{2\mu}(\boldsymbol{\gamma}, \operatorname{as} \boldsymbol{\tau}),$$

which, together with the Cauchy–Schwarz inequality, implies

$$\|\mathcal{A}\boldsymbol{\tau}\|^2 \le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}} + \|\nabla \mathbf{v}\|)\|\mathcal{A}\boldsymbol{\tau}\| + \frac{1}{2\mu}|(\boldsymbol{\gamma}, \operatorname{as} \boldsymbol{\tau})|.$$

By the ϵ-inequality, we have

$$\|\mathcal{A}\boldsymbol{\tau}\|^2 \le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0}) + \|\nabla \mathbf{v}\|^2 + |(\boldsymbol{\gamma}, \operatorname{as} \boldsymbol{\tau})|).$$

It follows from the triangle and the Cauchy–Schwarz inequalities and (3.5) that

$$\begin{aligned} |(\boldsymbol{\gamma}, \operatorname{as} \boldsymbol{\tau})| &= \Big|\frac{1}{2}\Big(2\boldsymbol{\gamma} + \frac{1}{2\mu} \operatorname{as} \boldsymbol{\tau} - \nabla \times \mathbf{v}, \operatorname{as} \boldsymbol{\tau}\Big) - \frac{1}{4\mu}\|\operatorname{as} \boldsymbol{\tau}\|^2 + \frac{1}{2}(\nabla \times \mathbf{v}, \operatorname{as} \boldsymbol{\tau})\Big| \\ &\le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0}) + \|\nabla \mathbf{v}\|^2). \end{aligned}$$

Hence,

$$\|\mathcal{A}\boldsymbol{\tau}\|^2 \le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0}) + \|\nabla \mathbf{v}\|^2),$$

which, together with (3.7), gives

$$\|\boldsymbol{\tau}\| \le C(F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})^{\frac{1}{2}} + \|\nabla \mathbf{v}\|). \tag{3.13}$$

Using (3.11), (3.12), (3.13), and the ϵ-inequality, we have

$$\|\nabla \mathbf{v}\|^2 + \|q\|^2 + \|\boldsymbol{\tau}\|^2 \le CF(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0}).$$

Now, the lower bound in (3.8) follows from the triangle and the Poincaré inequalities. This completes the proof of the theorem. □

4 Finite Element Approximation

Let us assume that Ω is a polygonal domain, let h be measure for the mesh-size according to [11] and let $\mathcal{T}_h = \{K\}$ be a finite element partition of Ω, which is regular (see [22]). For convenience, we consider only triangular and tetrahedral elements.

Since the least-square functionals $G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{0})$ and $F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{0})$ are equivalent to the energy norms $|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma})|||_{\mathcal{V}}$ and $|||(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q)|||_{\mathcal{W}}$, respectively, which are composed of the $H(\operatorname{div}; \Omega)$ norm for the stress, the $H^1(\Omega)$ norm for the displacement/velocity, and the $L^2(\Omega)$ norm for the rotation/vorticity and pressure by Theorem 1, it is reasonable to use Raviart–Thomas space of index k for the stress (see [26]), standard (conforming) continuous piecewise polynomials of degree $k + 1$ for the displacement/velocity, and the continuous/discontinuous piecewise polynomials of degree k for the rotation/vorticity and pressure as the conforming finite element spaces. These spaces are denoted as follows:

$$\begin{aligned} \Sigma_h^k &= \{\boldsymbol{\tau} \in \Sigma_N : \boldsymbol{\tau}|_K \in \mathrm{RT}_k(K)^d \text{ for all } K \in \mathcal{T}_h\} \subset \Sigma_N, \\ \mathcal{U}_h^k &= \{\mathbf{v} \in C^0(\Omega)^d : \mathbf{v}|_K \in P_k(K)^d \text{ for all } K \in \mathcal{T}_h,\ \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D\} \subset H_D^1(\Omega)^d, \\ \mathcal{L}_h^k &= \Big\{\boldsymbol{\gamma} \in L^2(\Omega) : \boldsymbol{\gamma}|_K \in P_k(K) \text{ for all } K \in \mathcal{T}_h,\ \int_\Omega \boldsymbol{\gamma}\, dx = 0 \text{ if } \Gamma_N = \emptyset\Big\} \subset L_D^2(\Omega), \end{aligned}$$

where $\mathrm{RT}_k(K)$ is local Raviart–Thomas space of index k defined by

$$\mathrm{RT}_k(K) = P_k(K)^d + \mathbf{x}P_k(K),$$

and $P_k(K)$ is the space of polynomials of degree k on K. They have the following approximation properties [13]: let $k \geq 0$ be an integer and let $l \in (0, k+1]$; then

$$\inf_{\boldsymbol{\tau} \in \Sigma_h^k} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{H(\mathrm{div};\Omega)} \leq Ch^l(\|\boldsymbol{\sigma}\|_l + \|\nabla \cdot \boldsymbol{\sigma}\|_l) \tag{4.1}$$

for $\boldsymbol{\sigma} \in H^l(\Omega)^{d\times d} \cap \Sigma_N$ with $\nabla \cdot \boldsymbol{\sigma} \in H^l(\Omega)^d$,

$$\inf_{\mathbf{v} \in \mathcal{U}_h^{k+1}} \|\mathbf{u} - \mathbf{v}\|_1 \leq Ch^l\|\mathbf{u}\|_{l+1} \tag{4.2}$$

for $\mathbf{u} \in H^{l+1}(\Omega)^d \cap H_D^1(\Omega)^d$, and

$$\inf_{\boldsymbol{\gamma} \in \mathcal{L}_h^k} \|\boldsymbol{\omega} - \boldsymbol{\gamma}\| \leq Ch^l\|\boldsymbol{\omega}\|_l \tag{4.3}$$

for $\boldsymbol{\omega} \in H^l(\Omega)^{2d-3} \cap L_D^2(\Omega)^{2d-3}$.

The finite element approximation for the least-squares problem in (3.2) on $\mathcal{V}_h^k = \Sigma_h^k \times \mathcal{U}_h^{k+1} \times (\mathcal{L}_h^k)^{2d-3}$ is defined as follows: find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h) \in \mathcal{V}_h^k$ such that

$$G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h; \mathbf{f}) = \min_{(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma}) \in \mathcal{V}_h^k} G(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}; \mathbf{f}). \tag{4.4}$$

By the fact that $\mathcal{V}_h^k \subset \mathcal{V}$, Theorem 1 implies that (4.4) has a unique solution and is equivalent to the weak form: find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h) \in \mathcal{V}_h^k$ such that

$$a(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) = -(\mathbf{f}, \nabla \cdot \boldsymbol{\tau}) \quad \text{for all } (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) \in \mathcal{V}_h^k,$$

where the bilinear form a is defined by

$$\begin{aligned} a(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) = {} & (\mathcal{A}_\lambda \boldsymbol{\sigma}_h - \nabla \mathbf{u}_h + (-1)^d \boldsymbol{\omega}_h \cdot \boldsymbol{\chi}, \mathcal{A}_\lambda \boldsymbol{\tau} - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi}) \\ & + (\nabla \cdot \boldsymbol{\sigma}_h, \nabla \cdot \boldsymbol{\tau}) + (\mathrm{as}\, \boldsymbol{\sigma}_h, \mathrm{as}\, \boldsymbol{\tau}). \end{aligned}$$

Similarly, the minimization in (3.3) on the space $\mathcal{W}_h^k = \Sigma_h^k \times \mathcal{U}_h^{k+1} \times (\mathcal{L}_h^k)^{2d-3} \times \mathcal{L}_h^k$ is defined as follows: find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h) \in \mathcal{W}_h^k$ such that

$$F(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h; \mathbf{f}) = \min_{(\boldsymbol{\tau},\mathbf{v},\boldsymbol{\gamma},q) \in \mathcal{W}_h^k} F(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q; \mathbf{f}). \tag{4.5}$$

Note that (4.5) has a unique solution and is equivalent to the weak form: find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h) \in \mathcal{W}_h^k$ such that

$$b(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) = -(\mathbf{f}, \nabla \cdot \boldsymbol{\tau}), \quad \forall (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) \in \mathcal{W}_h^k,$$

where the bilinear form b is defined by

$$\begin{aligned} & b(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h, ; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) \\ & = \Big(\frac{1}{2\mu}\Big(\boldsymbol{\sigma}_h + \frac{d\lambda}{2\mu + d\lambda} p_h \boldsymbol{\delta}\Big) - \nabla \mathbf{u}_h + (-1)^d \boldsymbol{\omega}_h \cdot \boldsymbol{\chi}, \frac{1}{2\mu}\Big(\boldsymbol{\tau} + \frac{d\lambda}{2\mu + d\lambda} q \boldsymbol{\delta}\Big) - \nabla \mathbf{v} + (-1)^d \boldsymbol{\gamma} \cdot \boldsymbol{\chi}\Big) \\ & \quad + (\nabla \cdot \boldsymbol{\sigma}_h, \nabla \cdot \boldsymbol{\tau}) + (\mathrm{as}\, \boldsymbol{\sigma}_h, \mathrm{as}\, \boldsymbol{\tau}) + \Big(\frac{d}{2\mu + d\lambda} p_h + \nabla \cdot \mathbf{u}_h, \frac{d}{2\mu + d\lambda} q + \nabla \cdot \mathbf{v}\Big). \end{aligned}$$

Furthermore, we have the following orthogonalities:

$$\begin{aligned} a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h, \boldsymbol{\omega} - \boldsymbol{\omega}_h; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) = 0 & \quad \text{for all } (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) \in \mathcal{V}_h^k, \\ b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h, \boldsymbol{\omega} - \boldsymbol{\omega}_h, p - p_h, ; \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) = 0 & \quad \text{for all } (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}, q) \in \mathcal{W}_h^k. \end{aligned} \tag{4.6}$$

Theorem 2 (Energy Norm Error Estimate). *Let $k+1$ be the smallest integer greater than or equal to $l > 0$.*

(i) *Assume that the solution $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\omega})$ of (3.2) belongs to $H^l(\Omega)^{d\times d} \times H^{l+1}(\Omega)^d \times H^l(\Omega)^{2d-3}$ with the divergence of the stress $\nabla\cdot\boldsymbol{\sigma}$ in $H^l(\Omega)^d$. Then, with the least-squares finite element solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h) \in \mathcal{V}_h^k$ of (4.4), we have the following error estimate:*

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(\mathrm{div};\Omega)} + \|\mathbf{u}-\mathbf{u}_h\|_1 + \|\boldsymbol{\omega}-\boldsymbol{\omega}_h\| \le Ch^l(\|\boldsymbol{\sigma}\|_l + \|\nabla\cdot\boldsymbol{\sigma}\|_l + \|\mathbf{u}\|_{l+1} + \|\boldsymbol{\omega}\|_l). \tag{4.7}$$

(ii) *Assume that the solution $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\omega}, p)$ of problem (3.3) belongs to $H^l(\Omega)^{d\times d} \times H^{l+1}(\Omega)^d \times H^l(\Omega)^{2d-3} \times H^l(\Omega)$ with the divergence of the stress $\nabla\cdot\boldsymbol{\sigma}$ in $H^l(\Omega)^d$. Then, with the least-squares finite element solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h) \in \mathcal{W}_h^k$ of (4.5), we have the following error estimate:*

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(\mathrm{div};\Omega)} + \|\mathbf{u}-\mathbf{u}_h\|_1 + \|\boldsymbol{\omega}-\boldsymbol{\omega}_h\| + \|p-p_h\| \le Ch^l(\|\boldsymbol{\sigma}\|_l + \|\nabla\cdot\boldsymbol{\sigma}\|_l + \|\mathbf{u}\|_{l+1} + \|\boldsymbol{\omega}\|_l + \|p\|_l). \tag{4.8}$$

Proof. The proof of the theorem follows directly from the coercivities and continuities of the bilinear forms in Theorem 1, the orthogonalities in (4.6), and the approximation properties in (4.1)–(4.3). □

In the remainder of this section, we establish optimal L^2-norm estimate of the error $\mathbf{u}-\mathbf{u}_h$.

Lemma 4. *Let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\omega})$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h)$ be the solutions of problem (3.2) and problem (4.4), respectively. Assume that the H^2 regularity estimate for the generalized Stokes equations (see [14]) holds. Then there exists $(\boldsymbol{\zeta}, \mathbf{w}, \boldsymbol{\varepsilon}) \in \Sigma_N \times H_D^1(\Omega)^d \times L^2(\Omega)^{2d-3}$ such that*

$$\|\mathbf{u}-\mathbf{u}_h\|^2 = a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h; \boldsymbol{\zeta}, \mathbf{w}, \boldsymbol{\varepsilon})$$

and that

$$\|\boldsymbol{\zeta}\|_1 + \|\nabla\cdot\boldsymbol{\zeta}\|_1 + \|\mathbf{w}\|_2 + \|\boldsymbol{\varepsilon}\|_1 \le C\|\mathbf{u}-\mathbf{u}_h\|.$$

Proof. Let $\mathbf{E}_h = \boldsymbol{\sigma}-\boldsymbol{\sigma}_h$, $\mathbf{e}_h = \mathbf{u}-\mathbf{u}_h$, and $\tilde{\mathbf{e}}_h = \boldsymbol{\omega}-\boldsymbol{\omega}_h$. Then, it is easy to check that

$$\begin{cases} \mathbf{n}\cdot\mathbf{E}_h = 0 & \text{on } \Gamma_N, \\ \mathbf{e}_h = 0 & \text{on } \Gamma_D. \end{cases}$$

Without loss of generality, take $\mu = \frac{1}{2}$. Let $(\mathbf{z}, r) \in H_D^1(\Omega)^d \times L^2(\Omega)$ be the solution of the following perturbed Stokes equation:

$$\begin{cases} -\nabla\cdot(\boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) = \mathbf{e}_h & \text{in } \Omega, \\ \nabla\cdot\mathbf{z} + \frac{1}{\lambda}r = 0 & \text{in } \Omega, \end{cases} \tag{4.9}$$

with the boundary conditions

$$\begin{cases} \mathbf{z} = \mathbf{0} & \text{on } \Gamma_D, \\ \mathbf{n}\cdot(\boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) = \mathbf{0} & \text{on } \Gamma_N. \end{cases}$$

It is well known that the following H^2 regularity estimate is valid:

$$\|\mathbf{z}\|_2 + \|r\|_1 \le C\|\mathbf{e}_h\|. \tag{4.10}$$

First, we derive the following equality:

$$\|\mathbf{e}_h\|^2 = (\mathcal{A}_\lambda\mathbf{E}_h - \nabla\mathbf{e}_h + (-1)^d\tilde{\mathbf{e}}_h\cdot\boldsymbol{\chi}, -\boldsymbol{\epsilon}(\mathbf{z}) + r\boldsymbol{\delta}) + (\nabla\cdot\mathbf{E}_h, -\mathbf{z}) + \Big(\operatorname{as}\mathbf{E}_h, -\frac{1}{2}\nabla\times\mathbf{z}\Big). \tag{4.11}$$

To this end, note first that

$$(\tilde{\mathbf{e}}_h\cdot\boldsymbol{\chi}, \boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) = 0.$$

Then the first equation in (4.9) and integration by parts give

$$\begin{aligned} \|\mathbf{e}_h\|^2 &= (\mathbf{e}_h, -\nabla\cdot(\boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta})) = (\nabla\mathbf{e}_h, \boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) \\ &= (\nabla\mathbf{e}_h - (-1)^d\tilde{\mathbf{e}}_h\cdot\boldsymbol{\chi} - \mathcal{A}_\lambda\mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) + (\mathcal{A}_\lambda\mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}). \end{aligned}$$

Using the definition of $\mathcal{A}_\lambda$ in (2.2), the facts that

$$\operatorname{tr}(\mathcal{A}_\lambda \mathbf{E}_h) = \frac{1}{d\lambda + 1} \operatorname{tr} \mathbf{E}_h \quad \text{and} \quad \boldsymbol{\delta} : \boldsymbol{\epsilon}(\mathbf{z}) = \nabla \cdot \mathbf{z},$$

and the second equation in (4.9), we have

$$\begin{aligned}(\mathcal{A}_\lambda \mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z}) - r\boldsymbol{\delta}) &= \Big(\mathbf{E}_h - \frac{\lambda}{d\lambda + 1}(\operatorname{tr} \mathbf{E}_h)\boldsymbol{\delta}, \boldsymbol{\epsilon}(\mathbf{z})\Big) - (\operatorname{tr}(\mathcal{A}_\lambda \mathbf{E}_h), r)\\ &= (\mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z})) - \frac{\lambda}{d\lambda + 1}(\operatorname{tr} \mathbf{E}_h, \nabla \cdot \mathbf{z}) - \frac{1}{d\lambda + 1}(\operatorname{tr} \mathbf{E}_h, r)\\ &= (\mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z})).\end{aligned}$$

Now, (4.11) follows from the above equalities, (2.5), integration by parts, and (2.6) that

$$\begin{aligned}(\mathbf{E}_h, \boldsymbol{\epsilon}(\mathbf{z})) &= (\mathbf{E}_h, \nabla \mathbf{z}) - \Big(\mathbf{E}_h, \frac{(-1)^d}{2}(\nabla \times \mathbf{z}) \cdot \boldsymbol{\chi}\Big)\\ &= (\nabla \cdot \mathbf{E}_h, -\mathbf{z}) + \Big(\operatorname{as} \mathbf{E}_h, -\frac{1}{2}\nabla \times \mathbf{z}\Big).\end{aligned}$$

With equation (4.11), to show the validity of Lemma 4, it is easy to see that it is sufficient to find $(\boldsymbol{\zeta}, \mathbf{w}, \boldsymbol{\varepsilon}) \in \Sigma_N \times H_D^1(\Omega)^d \times L^2(\Omega)^{2d-3}$ such that

$$\begin{cases} \mathcal{A}_\lambda \boldsymbol{\zeta} - \nabla \mathbf{w} + (-1)^d \boldsymbol{\varepsilon} \cdot \boldsymbol{\chi} = -\boldsymbol{\epsilon}(\mathbf{z}) + r\boldsymbol{\delta} & \text{in } \Omega,\\ \nabla \cdot \boldsymbol{\zeta} = -\mathbf{z} & \text{in } \Omega,\\ \operatorname{as} \boldsymbol{\zeta} = -\frac{1}{2}\nabla \times \mathbf{z} & \text{in } \Omega,\end{cases} \tag{4.12}$$

and that

$$\|\boldsymbol{\zeta}\|_1 + \|\nabla \cdot \boldsymbol{\zeta}\|_1 + \|\mathbf{w}\|_2 + \|\boldsymbol{\varepsilon}\|_1 \le C\|\mathbf{e}_h\|. \tag{4.13}$$

To this end, let $(\mathbf{w}, t) \in H_D^1(\Omega)^d \times L^2(\Omega)$ be the solution of the following problem:

$$\begin{cases} -\nabla \cdot (\boldsymbol{\epsilon}(\mathbf{w}) - t\boldsymbol{\delta}) = -\nabla \cdot \Big(\nabla \mathbf{z} - \frac{(-1)^d}{4}(\nabla \times \mathbf{z}) \cdot \boldsymbol{\chi}\Big) + \mathbf{z} & \text{in } \Omega,\\ \nabla \cdot \mathbf{w} + \frac{1}{\lambda} t = -\frac{d\lambda + 2}{\lambda} r & \text{in } \Omega,\end{cases} \tag{4.14}$$

with the boundary conditions

$$\begin{cases} \mathbf{w} = \mathbf{0} & \text{on } \Gamma_D,\\ \mathbf{n} \cdot (\boldsymbol{\epsilon}(\mathbf{w}) - t\boldsymbol{\delta}) = \mathbf{0} & \text{on } \Gamma_N.\end{cases}$$

It follows from the H^2 regularity estimate (see, e.g., [14, Lemma 5.2]), the triangle inequality, and (4.10) that

$$\begin{aligned}\|\mathbf{w}\|_2 + \|t\|_1 &\le C\Big(\Big\|-\nabla \cdot \Big(\nabla \mathbf{z} - \frac{(-1)^d}{4}(\nabla \times \mathbf{z}) \cdot \boldsymbol{\chi}\Big) + \mathbf{z}\Big\| + \Big\|\frac{d\lambda + 2}{\lambda} r\Big\|_1\Big)\\ &\le C(\|\mathbf{z}\|_2 + \|r\|_1)\\ &\le C\|\mathbf{e}_h\|.\end{aligned} \tag{4.15}$$

Let

$$\boldsymbol{\zeta} = -\nabla \mathbf{z} + \frac{(-1)^d}{4}(\nabla \times \mathbf{z}) \cdot \boldsymbol{\chi} + \boldsymbol{\epsilon}(\mathbf{w}) - t\boldsymbol{\delta}$$

and

$$\boldsymbol{\varepsilon} = \frac{1}{2}\nabla \times \mathbf{w} + \frac{1}{4}\nabla \times \mathbf{z}.$$

It is then easy to check that the second equation in (4.12) is satisfied by using the first equation in (4.14). The third equation in (4.12) is implied by (2.6) and the fact that

$$\operatorname{as}(\boldsymbol{\epsilon}(\mathbf{w}) - t\boldsymbol{\delta}) = \mathbf{0}.$$

From the second equations in (4.9) and (4.14), we have

$$\operatorname{tr}\boldsymbol{\zeta} = -\frac{d\lambda+1}{\lambda}(r+t). \tag{4.16}$$

By using (4.16) and the fact that

$$\boldsymbol{\epsilon}(\mathbf{w}) = \nabla\mathbf{w} - \frac{(-1)^d}{2}(\nabla\times\mathbf{w})\cdot\boldsymbol{\chi},$$

we can see that $(\boldsymbol{\zeta}, \mathbf{w}, \boldsymbol{\varepsilon})$ satisfies the first equation in (4.12).

It follows from the triangle inequality and the second equation in (4.12) that

$$\begin{aligned}\|\boldsymbol{\zeta}\|_1 + \|\nabla\cdot\boldsymbol{\zeta}\|_1 + \|\boldsymbol{\varepsilon}\|_1 &\le C(\|\mathbf{z}\|_2 + \|\mathbf{w}\|_2 + \|t\|_1) + \|\mathbf{z}\|_1 + C(\|\mathbf{w}\|_2 + \|\mathbf{z}\|_2)\\ &\le C(\|\mathbf{z}\|_2 + \|\mathbf{w}\|_2 + \|t\|_1),\end{aligned}$$

which, together with (4.10) and (4.15), implies the validity of (4.13). This completes the proof. □

Theorem 3 (L^2-Norm Error Estimate). *The following statements hold:*

(i) *Under assumption* (i) *in Theorem 2 and the assumptions in Lemma 4, we have the following L^2-norm error estimate:*

$$\|\mathbf{u}-\mathbf{u}_h\| \le Ch|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h)|||_{\mathcal{V}} \le Ch^{l+1}(\|\boldsymbol{\sigma}\|_l + \|\nabla\cdot\boldsymbol{\sigma}\|_l + \|\mathbf{u}\|_{l+1} + \|\boldsymbol{\omega}\|_l). \tag{4.17}$$

(ii) *Let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\omega}, p)$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h, p_h)$ be the solutions of* (3.3) *and* (4.5)*, respectively. Assume that the H^2 regularity estimate for the generalized Stokes equations (see* [14]*) holds. Then, under assumption* (ii) *in Theorem 2, we have the following L^2-norm error estimate:*

$$\|\mathbf{u}-\mathbf{u}_h\| \le Ch|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h, p-p_h)|||_{\mathcal{W}} \le Ch^{l+1}(\|\boldsymbol{\sigma}\|_l + \|\nabla\cdot\boldsymbol{\sigma}\|_l + \|\mathbf{u}\|_{l+1} + \|\boldsymbol{\omega}\|_l + \|p\|_l).$$

Proof. The second inequality in (4.17) is a direct consequence of the first inequality in (4.17) and the energy norm error estimate in (4.7). The first inequality in (4.17) follows from Lemma 4, the orthogonality, the continuity in (3.9), and the approximation properties in (4.1)–(4.3) that

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|^2 &= a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h; \boldsymbol{\zeta}, \mathbf{w}, \boldsymbol{\varepsilon})\\ &\le C|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h)|||_{\mathcal{V}} \inf_{(\boldsymbol{\zeta}_h, \mathbf{w}_h, \boldsymbol{\varepsilon}_h)\in\mathcal{V}_h^k} |||(\boldsymbol{\zeta}-\boldsymbol{\zeta}_h, \mathbf{w}-\mathbf{w}_h, \boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}_h)|||_{\mathcal{V}}\\ &\le Ch|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h)|||_{\mathcal{V}}(\|\boldsymbol{\zeta}\|_1 + \|\nabla\cdot\boldsymbol{\zeta}\|_1 + \|\mathbf{w}\|_2 + \|\boldsymbol{\varepsilon}\|_1)\\ &\le Ch|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \mathbf{u}-\mathbf{u}_h, \boldsymbol{\omega}-\boldsymbol{\omega}_h)|||_{\mathcal{V}}\|\mathbf{u}-\mathbf{u}_h\|.\end{aligned}$$

This completes the proof of (i). The proof of (ii) follows with a similar argument. □

5 Numerical Experiments

In this section, we consider two-dimensional Cook's membrane problem of linear elasticity (see [27]). The problem is given by the rotated trapezoidal geometry with bases of 16 and 44 units (l), and a height of 48 units (l) as in Figure 1, i.e.,

$$\Omega = \left\{(x, y) : 0 \le x \le 48,\ \frac{11}{12}x \le y \le \frac{1}{3} + 44\right\}.$$

They are no body forces ($\mathbf{f} = \mathbf{0}$), the left side of the computational domain is clamped and the boundary conditions on it are set to $\mathbf{u} = \mathbf{0}$ (l). The boundary conditions on the upper and lower sides are set to $\mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{0}$ (N/l^2) and on the right side are set to $\mathbf{n}\cdot\boldsymbol{\sigma} = (0, 1)^t$ (N/l^2) (see Table 1). The material parameters are $E = 200$ (N/l^2) for Young's modulus and $\nu = 0.499$ (incompressible limit) for Poisson's ratio and their relation with the Lamé constants are given by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{and} \quad \mu = \frac{E}{2(1+\nu)}.$$

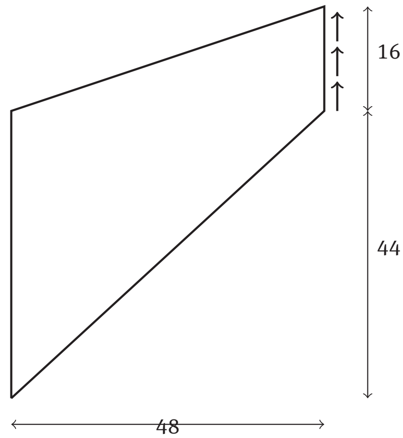


Figure 1: Computational domain and boundary conditions.

Boundary conditions		Material parameters	
Left side	$\mathbf{u} = \mathbf{0}$	Young's module	$E = 200$
Upper/Lower sides	$\mathbf{n} \cdot \boldsymbol{\sigma} = \mathbf{0}$	Poisson's ratio	$\nu = 0.499$
Right side	$\mathbf{n} \cdot \boldsymbol{\sigma} = (0, 1)^t$		

Table 1: Boundary conditions and material parameters.

	DOF	$G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h; \mathbf{f})^{\frac{1}{2}}$	$\Vert\mathrm{as}\,\boldsymbol{\sigma}_h\Vert$	$\Vert\nabla \cdot \boldsymbol{\sigma}_h + \mathbf{f}\Vert$
$l = 0$	1474	2.27e−01	7.93e−01	2.17e−01
$l = 1$	2594	2.99e−02	2.91e−01	2.88e−02
$l = 2$	4431	2.68e−03	8.37e−02	2.23e−03
$l = 3$	7513	3.53e−04	2.18e−02	1.47e−04
$l = 4$	12853	1.02e−04	5.50e−03	9.36e−06
$l = 5$	21572	4.28e−05	1.38e−03	5.90e−07
$l = 6$	37380	1.94e−05	3.45e−04	3.75e−08

Table 2: Adaptive refinements ($k = 1$, $\nu = 0.499$).

From the definition of the least-squares functional in (3.1), it is natural to define the local functional $G_K(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f})$ as follows:

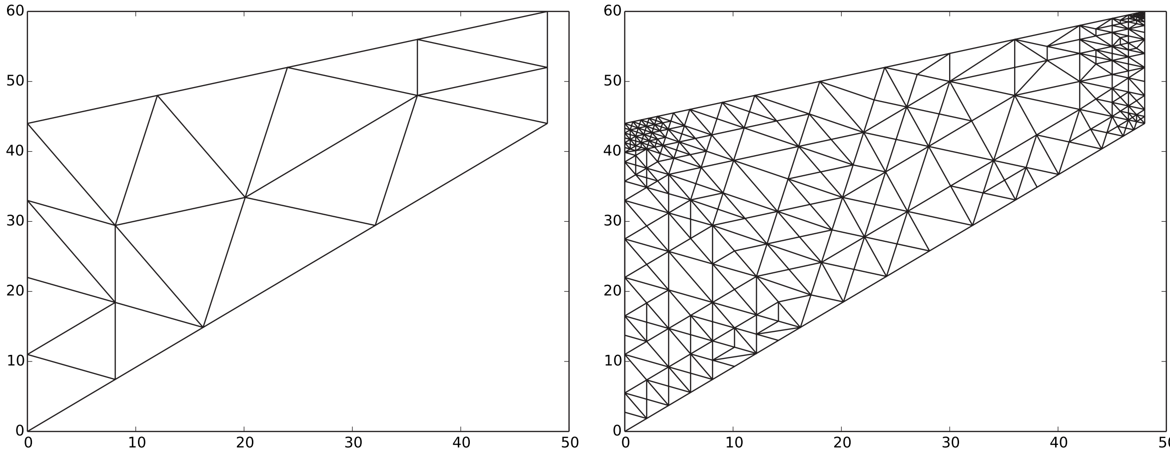
$$
\begin{aligned}
G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f}) &= \sum_{K \in \mathcal{T}_h} (\|\mathcal{A}_\lambda \boldsymbol{\sigma}_h - \nabla \mathbf{u}_h + \omega_h \boldsymbol{\chi}\|_{0,K}^2 + \|\nabla \cdot \boldsymbol{\sigma}_h + \mathbf{f}\|_{0,K}^2 + \|\mathrm{as}\,\boldsymbol{\sigma}_h\|_{0,K}^2) \\
&=: \sum_{K \in \mathcal{T}_h} G_K(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f}).
\end{aligned}
$$

By the equivalence in Theorem 1, the local functional $G_K(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\omega}_h; \mathbf{f})$ can be considered as a posteriori error estimator (see [6]). Table 2 shows the results obtained by a sequence of adaptive refinements based on this error estimator. In each step, triangles with the largest estimators (roughly 25 percent) were refined. We used the Raviart–Thomas spaces of order one for the stress approximation, the standard quadratic conforming elements for the displacement, and the standard linear conforming elements for the vorticity.

The result in Table 2 shows that the momentum balance error $\|\nabla \cdot \boldsymbol{\sigma}_h + \mathbf{f}\|$ is of higher order which implies the fast convergence (see Figure 3 (b)). We can also infer from Table 2 the following relation between the minimum of the functional $G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f})$ and the number of degrees of freedom:

$$
G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f}) \sim \frac{1}{\mathrm{DOF}^2},
$$

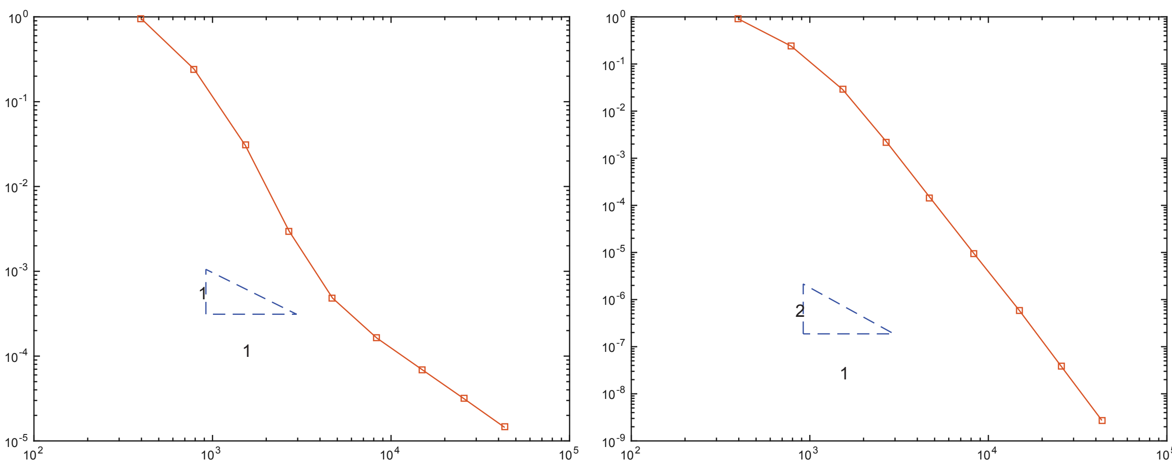
which implies the optimal asymptotic convergence rate.



(a) Initial mesh.

(b) Fifth adaptive refinement step.

Figure 2: Results after five adaptive refinement steps.



(a) $G(\boldsymbol{\sigma}_h, \mathbf{u}_h, \omega_h; \mathbf{f})$.

(b) $\|\nabla \cdot \boldsymbol{\sigma}_h + \mathbf{f}\|$.

Figure 3: Doubly logarithmic convergence graphs.

The initial triangulation and result after five adaptive refinement steps can be seen in Figure 2. It shows that the corners of the domain are remarkably refined well based on the error estimator. The doubly logarithmic convergence graphs in Figure 3 show the robustness of the problem. We can also see that the error converges asymptotically.

Funding: This work was supported in part by the National Science Foundation under grants DMS-1217081 and DMS-1522707.

References

[1] M. Amara and J. M. Thomas, Equilibrium finite elements for the linear elastic problem, *Numer. Math.* **33** (1979), no. 4, 367–383.

[2] D. N. Arnold, G. Awanou and R. Winther, Finite elements for symmetric tensors in three dimensions, *Math. Comp.* **77** (2008), no. 263, 1229–1251.

[3] D. N. Arnold, F. Brezzi and J. Douglas, Jr., PEERS: A new mixed finite element for plane elasticity, *Japan J. Appl. Math.* **1** (1984), no. 2, 347–367.

[4] D. N. Arnold, R. S. Falk and R. Winther, Mixed finite element methods for linear elasticity with weakly imposed symmetry, *Math. Comp.* **76** (2007), no. 260, 1699–1723.

[5] D. N. Arnold and R. Winther, Mixed finite elements for elasticity, *Numer. Math.* **92** (2002), no. 3, 401–419.

[6] M. Berndt, T. A. Manteuffel and S. F. McCormick, Local error estimates and adaptive refinement for first-order system least squares (FOSLS), *Electron. Trans. Numer. Anal.* **6** (1997), no. Dec., 35–43.

[7] P. B. Bochev and M. D. Gunzburger, Analysis of least squares finite element methods for the Stokes equations, *Math. Comp.* **63** (1994), no. 208, 479–506.

[8] P. B. Bochev and M. D. Gunzburger, Least-squares methods for the velocity-pressure-stress formulation of the Stokes equations, *Comput. Methods Appl. Mech. Engrg.* **126** (1995), no. 3–4, 267–287.

[9] P. B. Bochev and M. D. Gunzburger, Finite element methods of least-squares type, *SIAM Rev.* **40** (1998), no. 4, 789–837.

[10] P. B. Bochev and M. D. Gunzburger, *Least-Squares Finite Element Methods*, Appl. Math. Sci. 166, Springer, New York, 2009.

[11] D. Boffi, F. Brezzi and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Ser. Comput. Math. 44, Springer, Heidelberg, 2013.

[12] J. H. Bramble and J. E. Pasciak, Least-squares methods for Stokes equations based on a discrete minus one inner product, *J. Comput. Appl. Math.* **74** (1996), no. 1–2, 155–173.

[13] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer, New York, 1991.

[14] Z. Cai and J. Ku, The L^2 norm error estimates for the div least-squares method, *SIAM J. Numer. Anal.* **44** (2006), no. 4, 1721–1734.

[15] Z. Cai, B. Lee and P. Wang, Least-squares methods for incompressible Newtonian fluid flow: Linear stationary problems, *SIAM J. Numer. Anal.* **42** (2004), no. 2, 843–859.

[16] Z. Cai, T. A. Manteuffel and S. F. McCormick, First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity, *Electron. Trans. Numer. Anal.* **3** (1995), 150–159.

[17] Z. Cai, T. A. Manteuffel and S. F. McCormick, First-order system least squares for the Stokes equations, with application to linear elasticity, *SIAM J. Numer. Anal.* **34** (1997), no. 5, 1727–1741.

[18] Z. Cai, T. A. Manteuffel, S. F. McCormick and S. V. Parter, First-order system least squares (FOSLS) for planar linear elasticity: Pure traction problem, *SIAM J. Numer. Anal.* **35** (1998), no. 1, 320–335.

[19] Z. Cai and G. Starke, First-order system least squares for the stress-displacement formulation: Linear elasticity, *SIAM J. Numer. Anal.* **41** (2003), no. 2, 715–730.

[20] Z. Cai and G. Starke, Least-squares methods for linear elasticity, *SIAM J. Numer. Anal.* **42** (2004), no. 2, 826–842.

[21] C. L. Chang, A mixed finite element method for the Stokes problem: An acceleration-pressure formulation, *Appl. Math. Comput.* **36** (1990), no. 2, 135–146.

[22] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2002.

[23] B. M. Fraeijs de Veubeke, Stress function approach, in: *Proceedings of the World Congress on Finite Element Methods in Structural Mechanics. Volume 1*, Bournemouth (1975), 12–17.

[24] B.-N. Jiang, *The Least-Squares Finite Element Method. Theory and Applications in Computational Fluid Dynamics and Electromagnetics*, Sci. Comput., Springer, Berlin, 1998.

[25] B.-N. Jiang and J. Wu, The least-squares finite element method in elasticity. I. Plane stress or strain with drilling degrees of freedom, *Internat. J. Numer. Methods Engrg.* **53** (2002), no. 3, 621–636.

[26] P.-A. Raviart and J.-M. Thomas, A mixed finite element method for 2-nd order elliptic problems, in: *Mathematical Aspects of Finite Element Methods*, Springer, Berlin (1977), 292–315.

[27] A. Schwarz, J. Schröder and G. Starke, A modified least-squares mixed finite element with improved momentum balance, *Internat. J. Numer. Methods Engrg.* **81** (2010), no. 3, 286–306.

[28] G. Starke, A. Schwarz and J. Schröder, Analysis of a modified first-order system least squares method for linear elasticity with improved momentum balance, *SIAM J. Numer. Anal.* **49** (2011), no. 3, 1006–1022.