Journal of Computational Mathematics, Vol.21, No.6, 2003, 689–702.

DISCRETE MINUS ONE NORM LEAST-SQUARES FOR THE STRESS FORMULATION OF LINEAR ELASTICITY WITH NUMERICAL RESULTS *1)

Sang Dong Kim
(*Department of Mathematics, Kyungpook National University, Taegu 702-701, Korea*)

Byeong Chun Shin
(*Department of Mathematics, Chonnam National University, Kwangju 500-757, Korea*)

Seokchan Kim Gyungsoo Woo
(*Department of Mathematics, Changwon Nationnal University, Changwon 641-773, Korea*)

Abstract

This paper studies the discrete minus one norm least-squares methods for the stress formulation of pure displacement linear elasticity in two dimensions. The proposed least-squares functional is defined as the sum of the L^2- and H^{-1}-norms of the residual equations weighted appropriately. The minus one norm in the functional is replaced by the discrete minus one norm and then the discrete minus one norm least-squares methods are analyzed with various numerical results focusing on the finite element accuracy and multigrid convergence performances.

Key words: H^{-1} least-squares, Linear elasticity, Multigrid method.

1. Introduction

In recent years there has been an increased interest in the use of least-squares methods for numerical approximation of the incompressible Stokes and Navier-Stokes equations [3, 4, 5, 6, 7, 11, 12] and for linear elasticity equations [9, 10, 11, 12, 13, 18, 21]. Such least-squares approaches are known as including accurate approximations to meaningful physical quantities, formulation of a well-posed minimization principle and freedom in the choice of finite element spaces which are not subject to the LBB condition.

In this paper, we attempt to apply H^{-1} least-squares method to planar linear elasticity equations with pure displacement boundary conditions:

$$\begin{cases} -\mu\Delta\mathbf{u} - (\lambda+\mu)\nabla\nabla\cdot\mathbf{u} = \mathbf{f} & \text{in} \quad \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on} \quad \partial\Omega, \end{cases} \tag{1.1}$$

where Ω is a bounded open connected domain in $\Re^2$ with Lipschitz boundary $\partial\Omega$; $\mathbf{u}$ denotes the displacement; $\mathbf{f}$ is a given body force; and μ, $\lambda > 0$ are the Lamé constants. We assume that the elastic material is isotropic, homogeneous, and strongly elliptic. Denote by the Poisson ratio $\nu = \frac{\lambda}{2(\lambda+\mu)} \in (0, \frac{1}{2})$.

It is well known that standard Galerkin finite element formulations for elasticity problem using piecewise linear elements are accurate for moderate values of a Lamé constant λ, but, as the elastic material becomes nearly incompressible, i.e. as $\lambda \to \infty$ (or $\nu \to \frac{1}{2}$), their approximation properties degrade severely [1, 20]. To overcome this, so-called *locking phenomenon*, Cai, Manteuffel, McCormick and their coworkers have developed recently first-order system

* Received August 20, 2001, final revised June 7, 2002.
1) The paper was supported by Korea Research Foundation Grant (KRF-2002-070-C00014).

L^2-norm and H^{-1}-norm least squares methods with the flux and vorticity formulation for the generalized Stokes equations that apply to the pure displacement problem of linear elasticity in [11, 13], and for the pure traction problem in [12]. In our formulation, defining new two variables as the strain tensor $\underset{\sim}{\boldsymbol{\sigma}} = \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})$ scaled by $\sqrt{2\mu}$ and pressure $p = -\nabla \cdot \mathbf{u}$, the second-order pure displacement problem is reduced to first-order system of linear equations, so-called strain-displacement-pressure formulation. In the analysis of structural mechanics, the knowledge of the stress or strain is often of greater interest than the knowledge of the displacement. Even though the approximation of the stress or strain can be recovered from the displacement by postprocessing in the standard finite element formulation, in a numerical point of view their computations require the derivatives of the displacement which imply a loss of precision. But, in the strain-displacement-pressure formulation we used, the accurate strain can be obtained directly and the stress can be directly recovered as the linear combination $-\lambda p\,\mathrm{I}_2 + \sqrt{2\mu}\,\underset{\sim}{\boldsymbol{\sigma}}$ of strain tensor $\underset{\sim}{\boldsymbol{\sigma}}$ and pressure p where I_2 is 2×2 identity matrix. The similar formulation for the elasticity problem can be found in [21] and the stress formulation for the incompressible Stokes equations were applied to the mixed methods and stabilized Galerkin methods in [2, 15, 17].

Our least-squares functional is similar to that in [3] with $q = -1$ but it is appropriately weighted by a Lamé constant μ. We will directly establish ellipticity and continuity of the functional in a product norm involving Lamé constants μ and λ and the L^2- and H^1-norms. To make the computation of H^{-1}-norm to be feasible, we replace the H^{-1}-norm in the functional by the discrete H^{-1}-norm following the idea proposed by Bramble, Lazarov and Pasciak including discrete H^{-1}-norm least-squares approaches for scalar second-order elliptic equations in [6] and for the Stokes equations in [7]. Such discrete H^{-1} functional is shown to be uniformly equivalent to the Sobolev norms weighted by the Lamé constants. From this property we show that standard finite element discretization error estimates are optimal with respect to the order of approximation as well as the required regularity of the solution, and that they are uniform in the Lamé constants.

The paper is organized as follows. In section 2, we formulate an equivalent first-order systems with the strain-displacement-pressure formulation to pure elasticity problem and set some preliminary results. We introduce H^{-1}-norm least-squares functional weighted appropriately by Lamé constant μ for the strain formulation and then we establish its ellipticity and continuity in section 3. In section 4, we consider discrete H^{-1}-norm least-squares functional and discuss an error estimate according to [6] and [7]. Finally section 5 investigates a preconditioner for the resulting algebraic linear system and present the numerical results implemented by preconditioned Richardson iteration method and multigrid V-cycle algorithm using continuous piecewise linear finite element spaces.

2. First-Order System Formulations

In this section we formulate a first-order system for H^{-1} least-squares methods with the strain formulation that is equivalent to the system of equations of linear elasticity with pure displacement boundary conditions.

For convenience, we let the boldface denote the vector valued function and the under tilde boldface the matrix-valued function, i.e., the tensor. We use C with or without subscripts to denote a generic positive constant, possibly different at different occurrences, that is independent of the Lamé constants and other parameters introduced in this paper, but may depend on the domain Ω. The colon notation : denotes the inner product on $\Re^{2\times 2}$ and for any tensors $\underset{\sim}{\boldsymbol{\tau}} = (\tau_{ij})$ and $\underset{\sim}{\boldsymbol{\delta}} = (\delta_{ij})$ in $L^2(\Omega)^{2\times 2}$, the $L^2(\Omega)^{2\times 2}$ inner product is defined by

$$(\underset{\sim}{\boldsymbol{\tau}}, \underset{\sim}{\boldsymbol{\delta}}) = \int_\Omega \underset{\sim}{\boldsymbol{\tau}} : \underset{\sim}{\boldsymbol{\delta}}\, dx.$$

The divergence for a tensor $\underset{\sim}{\tau}$ is defined as

$$\nabla \cdot \underset{\sim}{\tau} = \left(\sum_{j=1}^{2} \frac{\partial \tau_{1j}}{\partial x_j}, \sum_{j=1}^{2} \frac{\partial \tau_{2j}}{\partial x_j} \right)^t$$

where t denotes transpose.

Let $\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})$ denote the strain tensor:

$$\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) := \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^t) = (\epsilon_{i,j}), \quad \epsilon_{i,j}(\mathbf{u}) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2.$$

The stress tensor $\underset{\sim}{\tau}$ is given by

$$\underset{\sim}{\tau} := \lambda \, (\nabla \cdot \mathbf{u}) \mathrm{I}_2 + 2\mu \, \underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) \tag{2.1}$$

where I_2 denotes the 2×2 identity matrix. Let $\underset{\sim}{\boldsymbol{\sigma}}$ be the strain tensor scaled by $\sqrt{2\mu}$

$$\underset{\sim}{\boldsymbol{\sigma}} = \sqrt{2\mu} \, \underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) \tag{2.2}$$

and let us introduce a new variable p such as

$$p = -\nabla \cdot \mathbf{u}. \tag{2.3}$$

Then, the stress tensor $\underset{\sim}{\tau}$ is defined as the sum of the new variables $\underset{\sim}{\boldsymbol{\sigma}}$ and p:

$$\underset{\sim}{\tau} = -\lambda \, p \, \mathrm{I}_2 + \sqrt{2\mu} \, \underset{\sim}{\boldsymbol{\sigma}}.$$

Using the vector identity

$$\nabla \cdot \underset{\sim}{\boldsymbol{\sigma}} = \frac{\sqrt{2\mu}}{2} (\Delta \mathbf{u} + \nabla \nabla \cdot \mathbf{u}),$$

we have the first-order system of the *strain formulation* which is equivalent to (1.1):

$$\left\{ \begin{array}{rcll} \sqrt{2\mu} \, \nabla \cdot \underset{\sim}{\boldsymbol{\sigma}} - \lambda \, \nabla p & = & -\mathbf{f} & \text{in} \quad \Omega, \\ \underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu} \, \underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) & = & \underset{\sim}{\mathbf{0}} & \text{in} \quad \Omega, \\ \nabla \cdot \mathbf{u} + p & = & 0 & \text{in} \quad \Omega, \\ \mathbf{u} & = & \mathbf{0} & \text{on} \quad \partial\Omega \end{array} \right. \tag{2.4}$$

where p satisfies the compatibility condition

$$\int_\Omega p \, dx = 0.$$

The system (2.4) can be reduced to a system of six unknowns and six equations because the second equations in (2.4) consists of only three equations in virtue of symmetric tensors $\underset{\sim}{\boldsymbol{\sigma}}$ and $\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})$.

In order to give proper formulation of H^{-1} least-squares methods we use the Sobolev spaces $H^s(\Omega)$ with the standard associated inner products $(\cdot, \cdot)_s$ and their respective norms $\| \cdot \|_s$. For $s = 0$, $H^s(\Omega)$ coincides with $L^2(\Omega)$. In this case the norm and inner product will be denoted by

$\|\cdot\|$ and $(\cdot,\cdot)$, respectively. As usual, $H_0^s(\Omega)$ will denote the closure of $\mathcal{D}(\Omega)$, the linear space of infinitely differentiable functions with compact supports on Ω, with respect to the norm $\|\cdot\|_s$ and $L_0^2(\Omega)$ will denote the subspace of square integrable functions with zero mean:

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

For positive values of s the space $H^{-s}(\Omega)$ is defined as the dual space of $H_0^s(\Omega)$ equipped with the norm

$$\|\phi\|_{-s} := \sup_{0 \neq v \in H_0^s(\Omega)} \frac{(\phi, v)}{\|v\|_s}.$$

Here, $(\cdot,\cdot)$ is also used as the duality pairing between $H^{-s}(\Omega)$ and $H_0^s(\Omega)$ when there is no risk of confusion. Define the product spaces $H_0^s(\Omega)^2 = \prod_{i=1}^2 H_0^s(\Omega)$ and $H^{-s}(\Omega)^2 = \prod_{i=1}^2 H^{-s}(\Omega)$ with standard product norms. Let

$$H(\mathrm{div};\Omega) = \{ \mathbf{v} \in L^2(\Omega)^2 \ : \ \nabla \cdot \mathbf{v} \in L^2(\Omega) \ \}$$

equipped with the norm

$$\|\mathbf{v}\|_{H(\mathrm{div};\Omega)} := \left(\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2 \right)^{1/2}.$$

We relate the particular space $H^{-1}(\Omega)^2$ to the boundary value problem as follows (see [6]): Consider the Dirichlet problem

$$\begin{cases} -\Delta \mathbf{w} + \mathbf{w} = \mathbf{f} & \text{in} \quad \Omega \\ \mathbf{w} = \mathbf{0} & \text{on} \quad \partial\Omega. \end{cases} \tag{2.5}$$

Let $T : H^{-1}(\Omega)^2 \to H_0^1(\Omega)^2$ denote the solution operator for the above problem (2.5), that is, for any $\mathbf{f} \in H^{-1}(\Omega)^2$, $T\mathbf{f} = \mathbf{w}$ is the solution to (2.5). From the definition of T, we can easily show that

$$\|\boldsymbol{\phi}\|_{-1}^2 = (\boldsymbol{\phi}, T\boldsymbol{\phi}). \tag{2.6}$$

Thus the inner product $(\cdot,\cdot)_{-1}$ on $H^{-1}(\Omega)^2 \times H^{-1}(\Omega)^2$ associated with the norm $\|\cdot\|_{-1}$ is given by

$$(\boldsymbol{\phi}, \boldsymbol{\psi})_{-1} = (\boldsymbol{\phi}, T\boldsymbol{\psi}) = (T\boldsymbol{\phi}, \boldsymbol{\psi}).$$

We now give the following useful identity

$$\int_\Omega \underset{\sim}{\epsilon}(\mathbf{u}) : \underset{\sim}{\tau} dx = -\int_\Omega (\nabla \cdot \underset{\sim}{\tau}) \cdot \mathbf{u} \, dx \tag{2.7}$$

for any symmetric tensor $\underset{\sim}{\tau}$ and any vector function $\mathbf{u}$ vanishing on the boundary, and finally let us recall from the consequences of a general functional analysis result in [19] and [16] that

$$\|p\| \leq C \|\nabla p\|_{-1} \quad \text{for any} \quad p \in L_0^2(\Omega). \tag{2.8}$$

3. Minus One Norm Least Squares Methods

In this section we define and analyze H^{-1} least-squares methods to approximate the solution of (2.4).

First, we take the least-squares functional as the sum of L^2- or H^{-1}-norms of the weighted residuals for the system (2.4) such that

$$G(\underset{\sim}{\sigma},\mathbf{u},p;\mathbf{f}) = \|\sqrt{2\mu}\,\nabla\cdot\underset{\sim}{\sigma} - \lambda\nabla p + \mathbf{f}\|_{-1}^2 + \mu\|\underset{\sim}{\sigma} - \sqrt{2\mu}\,\underset{\sim}{\epsilon}(\mathbf{u})\|^2 + \mu^2\|\nabla\cdot\mathbf{u} + p\|^2. \tag{3.1}$$

For a solution space, define

$$\boldsymbol{\mathcal{V}} := \tilde{\mathbf{L}}^2(\Omega)\times H_0^1(\Omega)^2\times L_0^2(\Omega)$$

where $\tilde{\mathbf{L}}^2(\Omega)$ is the space of 2×2 symmetric matrix functions whose elements are square-integrable.

The first-order system least-squares variational problem for the system (2.4) is to minimize the quadratic functional $G(\underset{\sim}{\sigma},\mathbf{u},p;\mathbf{f})$ over $\boldsymbol{\mathcal{V}}$: find $(\underset{\sim}{\sigma},\mathbf{u},p)\in\boldsymbol{\mathcal{V}}$ such that

$$G(\underset{\sim}{\sigma},\mathbf{u},p;\mathbf{f}) \;=\; \inf_{(\underset{\sim}{\tau},\mathbf{v},q)\in\boldsymbol{\mathcal{V}}} G(\underset{\sim}{\tau},\mathbf{v},q;\mathbf{f})\ . \tag{3.2}$$

Define a norm $|||\cdot|||$ on $\boldsymbol{\mathcal{V}}$ by

$$|||(\underset{\sim}{\tau},\mathbf{v},q)||| = \Big(\mu\|\underset{\sim}{\sigma}\|^2 + \mu^2\|\mathbf{u}\|_1^2 + \lambda^2\|p\|^2\Big)^{\frac{1}{2}}, \qquad \forall\,(\underset{\sim}{\tau},\mathbf{v},q)\in\boldsymbol{\mathcal{V}}.$$

From now on, we assume that $\mu\le\lambda$. Under this assumption, we show that the continuity and ellipticity of the functional $G(\underset{\sim}{\sigma},\mathbf{u},p;\mathbf{0})$ are independent of the Lamé constants λ and μ in the next theorem and the proposed functional is robust as the Poisson ratio $\nu\to\frac{1}{2}$.

Theorem 3.1. *For any $(\underset{\sim}{\sigma},\mathbf{u},p)\in\boldsymbol{\mathcal{V}}$, there exists a constant C, independent of μ and λ such that*

$$\frac{1}{C}\,|||(\underset{\sim}{\sigma},\mathbf{u},p)|||^2 \;\le\; G(\underset{\sim}{\sigma},\mathbf{u},p;\mathbf{0}) \;\le\; C\,|||(\underset{\sim}{\sigma},\mathbf{u},p)|||^2. \tag{3.3}$$

Proof. The upper bound of (3.3) is straight forward from the triangle and Cauchy-Schwarz inequalities and the assumption $\mu\le\lambda$. If we show the validity of the lower bound of (3.3) for $(\underset{\sim}{\sigma},\mathbf{u},p)\in\mathbf{W}$ where $\mathbf{W} := \big(\tilde{\mathbf{L}}^2(\Omega)\cap H(\mathrm{div};\Omega)^2\big)\times H_0^1(\Omega)^2\times\big(L_0^2(\Omega)\cap H^1(\Omega)\big)$, then the lower bound of (3.3) would follow for any $(\underset{\sim}{\sigma},\mathbf{u},p)\in\boldsymbol{\mathcal{V}}$ by continuity. For any $(\underset{\sim}{\sigma},\mathbf{u},p)\in\mathbf{W}$ and $\mathbf{v}\in H_0^1(\Omega)^2$, using the Schwarz inequality and the fact that $\|\underset{\sim}{\epsilon}(\mathbf{v})\|\le\|\mathbf{v}\|_1$ yield

$$\begin{aligned}
\big(\lambda\nabla p,\mathbf{v}\big) &= \Big|\big(-\sqrt{2\mu}\nabla\cdot\underset{\sim}{\sigma}+\lambda\nabla p,\mathbf{v}\big) + \big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\sigma},\mathbf{v}\big)\Big| \\
&= \Big|\big(-\sqrt{2\mu}\nabla\cdot\underset{\sim}{\sigma}+\lambda\nabla p,\mathbf{v}\big) - \big(\underset{\sim}{\sigma},\sqrt{2\mu}\underset{\sim}{\epsilon}(\mathbf{v})\big)\Big| \\
&= \Big|\big(-\sqrt{2\mu}\nabla\cdot\underset{\sim}{\sigma}+\lambda\nabla p,\mathbf{v}\big) - \sqrt{2\mu}\big(\underset{\sim}{\sigma}-\sqrt{2\mu}\underset{\sim}{\epsilon}(\mathbf{u}),\underset{\sim}{\epsilon}(\mathbf{v})\big) - 2\mu\big(\underset{\sim}{\epsilon}(\mathbf{u}),\underset{\sim}{\epsilon}(\mathbf{v})\big)\Big| \\
&\le \Big\{\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\sigma}-\lambda\nabla p\|_{-1} + \sqrt{2\mu}\|\underset{\sim}{\sigma}-\sqrt{2\mu}\underset{\sim}{\epsilon}(\mathbf{u})\| + 2\mu\|\underset{\sim}{\epsilon}(\mathbf{u})\|\Big\}\,\|\mathbf{v}\|_1.
\end{aligned}$$

Hence, from (2.8) we have

$$\|\lambda p\| \le \bar{C}\left\{\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \nabla p\|_{-1} + \sqrt{\mu}\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\| + \mu\|\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|\right\}. \tag{3.4}$$

Using (2.7) and the Schwarz inequality yields

$$\begin{aligned}
2\mu\|\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|^2 &= \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) - \underset{\sim}{\boldsymbol{\sigma}}, \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\big) + \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\sigma}}, \underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\big)\\
&= \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) - \underset{\sim}{\boldsymbol{\sigma}}, \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\big) - \big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}}, \mathbf{u}\big)\\
&= \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) - \underset{\sim}{\boldsymbol{\sigma}}, \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\big) - \big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p, \mathbf{u}\big) - \big(\lambda\nabla p, \mathbf{u}\big)\\
&= \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}) - \underset{\sim}{\boldsymbol{\sigma}}, \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\big) - \big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p, \mathbf{u}\big) + \big(\nabla\cdot\mathbf{u} + p, \lambda p\big) - (p, \lambda p)\\
&\le \sqrt{2\mu}\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|\,\|\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\| + \|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}\,\|\mathbf{u}\|_1 + \|\nabla\cdot\mathbf{u} + p\|\,\|\lambda p\|.
\end{aligned}$$

By multiplying μ and using Korn's inequality and (3.4), we have

$$\begin{aligned}
2\mu^2\|\mathbf{u}\|_1^2 &\le C\left(\mu\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}\cdot\|\mathbf{u}\|_1 + \mu\sqrt{2\mu}\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|\cdot\|\mathbf{u}\|_1\right)\\
&\quad + C\left(\mu\|\nabla\cdot\mathbf{u} + p\|\right)\cdot\left(\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1} + \sqrt{\mu}\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\| + \mu\|\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|\right)\\
&\le \frac{C^2}{2}\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}^2 + \frac{\mu^2}{2}\|\mathbf{u}\|_1 + C^2\mu\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|^2 + \frac{\mu^2}{2}\|\mathbf{u}\|_1\\
&\quad + \frac{C\bar{C}}{2}\left(2\mu^2\|\nabla\cdot\mathbf{u} + p\|^2 + \|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}^2 + \mu\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|^2\right)\\
&\quad + \frac{C^2\bar{C}^2}{2}\mu^2\|\nabla\cdot\mathbf{u} + p\|^2 + \frac{\mu^2}{2}\|\mathbf{u}\|_1^2
\end{aligned} \tag{3.5}$$

which implies

$$\mu^2\|\mathbf{u}\|_1^2 \le C\left\{\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}^2 + \mu\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|^2 + \mu^2\|\nabla\cdot\mathbf{u} + p\|^2\right\}. \tag{3.6}$$

Combining (3.6) with (3.4) yields

$$\lambda^2\|p\|^2 \le C\,G(\underset{\sim}{\boldsymbol{\sigma}}, \mathbf{u}, p, \mathbf{0}). \tag{3.7}$$

On the other hand, we have

$$\begin{aligned}
\mu\|\underset{\sim}{\boldsymbol{\sigma}}\|^2 &= \big(\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}), \mu\underset{\sim}{\boldsymbol{\sigma}}\big) + \big(\sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}), \mu\underset{\sim}{\boldsymbol{\sigma}}\big)\\
&= \big(\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}), \mu\underset{\sim}{\boldsymbol{\sigma}}\big) - \mu\big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}}, \mathbf{u}\big)\\
&= \big(\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}), \mu\underset{\sim}{\boldsymbol{\sigma}}\big) - \mu\big(\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p, \mathbf{u}\big) + \mu\big(\nabla\cdot\mathbf{u} + p, \lambda p\big) - \mu\big(p, \lambda p\big)\\
&\le \mu\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|\,\|\underset{\sim}{\boldsymbol{\sigma}}\| + \mu\|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}\,\|\mathbf{u}\|_1 + \mu\|\nabla\cdot\mathbf{u} + p\|\,\|\lambda p\|\\
&\le \frac{1}{2}\Big(\mu\|\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\|^2 + \mu\|\underset{\sim}{\boldsymbol{\sigma}}\|^2 + \|\sqrt{2\mu}\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\nabla p\|_{-1}^2\\
&\qquad + \mu^2\|\mathbf{u}\|_1^2 + \mu^2\|\nabla\cdot\mathbf{u} + p\|^2 + \|\lambda p\|^2\Big),
\end{aligned}$$

so that, from (3.6) and (3.7), we are led to

$$\mu\|\underset{\sim}{\sigma}\|^2 \le C\, G(\underset{\sim}{\sigma}, \mathbf{u}, p; \mathbf{0}). \tag{3.8}$$

Hence the lower bound of (3.3) follows from (3.6), (3.7) and (3.8).

4. Discrete Minus One Norm Least-Squares Methods

In this section we will present the discrete H^{-1} least-squares finite element approximation, the well-posedness of the discrete problem and then establish optimal order error estimates for each variable.

For the finite element approximation, we assume that the domain Ω is a convex polygonal and that $\mathcal{T}_h$ is a partition of Ω into finite elements $\Omega = \cup_{K\in\mathcal{T}_h} K$ with $h = \max\{\mathrm{diam}(K) : K \in \mathcal{T}_h\}$. Assume that the triangulation $\mathcal{T}_h$ is regular and satisfies the inverse assumption (see [14]). Let $\boldsymbol{\mathcal{V}}^h := \underset{\sim}{\Phi}^h \times \mathcal{U}^h \times P^h$ be a finite element subspace of $\boldsymbol{\mathcal{V}}$ with the following properties: there exist a constant C and an integer s such that for all $\left(\underset{\sim}{\sigma}, \mathbf{u}, p\right) \in \left(\tilde{\mathbf{L}}^2(\Omega) \cap H^r(\Omega)^{d\times d}\right) \times \left(H_0^1(\Omega)^2 \cap H^{r+1}(\Omega)^2\right) \times \left(L_0^2(\Omega) \cap H^r(\Omega)\right)$, $1 \le r \le s$, there exists a triplet $(\underset{\sim}{\sigma}^I, \mathbf{u}^I, p^I) \in \boldsymbol{\mathcal{V}}^h$ such that

$$\inf_{\underset{\sim}{\sigma}^I \in \underset{\sim}{\Phi}^h} \left(\|\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^I\| + h\|\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^I\|_1 \right) \le C h^r \|\underset{\sim}{\sigma}\|_r, \tag{4.1}$$

$$\inf_{\mathbf{u}^I \in \mathcal{U}^h} \left(\|\mathbf{u} - \mathbf{u}^I\| + h\|\mathbf{u} - \mathbf{u}^I\|_1 \right) \le C h^{r+1} \|\mathbf{u}\|_{r+1} \tag{4.2}$$

and

$$\inf_{p^I \in P^h} \left(\|p - p^I\| + h\|p - p^I\|_1 \right) \le C h^r \|p\|_r. \tag{4.3}$$

Note that typical finite element spaces consisting of continuous piecewise polynomials with respect to quasi-uniform triangulations satisfy (4.1)-(4.3) (for more detail see [14]).

Let $(\underset{\sim}{\sigma}, \mathbf{u}, p)$ be the solution of (2.4) which obviously minimizes the functional (3.1) and let $(\underset{\sim}{\sigma}^h, \mathbf{u}^h, p^h)$ minimize the functional (3.1) over $\boldsymbol{\mathcal{V}}^h$. Denote by $\mathbf{e}_\sigma = \underset{\sim}{\sigma} - \underset{\sim}{\sigma}^h$, $\mathbf{e}_u = \mathbf{u} - \mathbf{u}^h$ and $\mathbf{e}_p = p - p^h$. If we use continuous piecewise polynomials of degree r for the variable $\mathbf{u}^h$ and piecewise polynomials of degree $r-1$ for the variables p^h and $\underset{\sim}{\sigma}^h$ where $r \ge 2$, then we can show from Theorem 3.1 and the approximation properties (4.1)-(4.3) that

$$|||(\mathbf{e}_\sigma, \mathbf{e}_u, \mathbf{e}_p)||| \le C\, h^r \left\{ \sqrt{\mu}\, \|\underset{\sim}{\sigma}\|_r + \mu\, \|\mathbf{u}\|_{r+1} + \lambda\, \|p\|_r \right\},$$

which is optimal with respect to the finite element functions used. The use of continuous piecewise polynomials of degree r ($r \ge 1$) for all unknowns yields the error estimate $O(h^r)$, but it does not give optimal error estimates for the variables p^h and $\underset{\sim}{\sigma}^h$. However, the use of a single approximating space for all variables simplifies programming of least-squares finite element methods.

Following the ideas suggested in [6, 7], we now replace the operator T by an equivalent and computable operator T_h. Let $\tilde{B}_h : H^{-1}(\Omega)^2 \to \mathcal{U}^h$ be the discrete solution operator $\mathbf{w} = \tilde{B}_h \boldsymbol{\psi} \in \mathcal{U}^h$ for the Dirichlet problem (2.5) defined by

$$(\nabla \mathbf{w}, \nabla \mathbf{v}) + (\mathbf{w}, \mathbf{v}) = (\boldsymbol{\psi}, \mathbf{v}), \qquad \forall\, \mathbf{v} \in \mathcal{U}^h. \tag{4.4}$$

Assume that there is a preconditioner $B_h : H^{-1}(\Omega)^2 \to \mathcal{U}^h$ for $\tilde{B}_h$ that is symmetric positive definite operator with respective to the $L^2(\Omega)^2$-inner product and spectrally equivalent to $\tilde{B}_h$, i.e., there exists a positive constant C not depending on the mesh size h such that

$$\frac{1}{C}(B_h\mathbf{v},\mathbf{v}) \le (\tilde{B}_h\mathbf{v},\mathbf{v}) \le C(B_h\mathbf{v},\mathbf{v}), \quad \mathbf{v}\in\mathcal{U}^h. \tag{4.5}$$

Define $T_h = \alpha\, h^2 I + \beta\, B_h$ for fixed positive constants α and β where I denotes the identity operator on $\mathcal{U}^h$ and the parameters α and β could be used to tune the iterative convergence rate. The order of convergence is not changed (see [6]). For convenience we assume that $\alpha = \beta = 1$ in this section.

Now, we define the discrete counterparts of the least-squares functional G as follows:

$$\begin{aligned} G_h(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p;\mathbf{f}) = &\Big(T_h(\sqrt{2\mu}\,\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\,\nabla p + \mathbf{f}), \sqrt{2\mu}\,\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}} - \lambda\,\nabla p + \mathbf{f}\Big) \\ &+ \mu\Big(\underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\,\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u}), \underset{\sim}{\boldsymbol{\sigma}} - \sqrt{2\mu}\,\underset{\sim}{\boldsymbol{\epsilon}}(\mathbf{u})\Big) + \mu^2\Big(\nabla\cdot\mathbf{u} + p, \nabla\cdot\mathbf{u} + p\Big). \end{aligned} \tag{4.6}$$

The purpose of the remainder of this section is to analyze least-squares approximation based on the above functional. The finite element approximation to (3.2) becomes: find $(\underset{\sim}{\boldsymbol{\sigma}}^h,\mathbf{u}^h,p^h)\in\boldsymbol{\mathcal{V}}^h$ such that

$$G_h(\underset{\sim}{\boldsymbol{\sigma}}^h,\mathbf{u}^h,p^h;\mathbf{f}) \;=\; \inf_{(\underset{\sim}{\boldsymbol{\tau}},\mathbf{v},q)\in\boldsymbol{\mathcal{V}}^h} G_h(\underset{\sim}{\boldsymbol{\tau}},\mathbf{v},q;\mathbf{f})\;. \tag{4.7}$$

Let $Q_h : L^2(\Omega)^2 \to \mathcal{U}^h$ denote the L^2-orthogonal projection operator onto $\mathcal{U}^h$ which satisfies

$$\|Q_h\mathbf{v}\|_1 \le C\,\|\mathbf{v}\|_1, \quad \forall\,\mathbf{v}\in H^1(\Omega)^2. \tag{4.8}$$

Then, the symmetry of B_h with respect to the $L^2(\Omega)^2$-inner product yields that $B_h = B_hQ_h$. Similarly, $\tilde{B}_h = \tilde{B}_hQ_h$. Thus, the spectral equivalence in (4.5) holds for every $\mathbf{v}\in L^2(\Omega)^2$. We recall from [6] that

$$\|(I-Q_h)\mathbf{v}\|_{-1} \;\le\; C\,h\,\|\mathbf{v}\|, \quad \forall\,\mathbf{v}\in L^2(\Omega)^2, \tag{4.9}$$

and

$$\|Q_h\mathbf{v}\|_{-1}^2 \;\le C\;(\tilde{B}_h\mathbf{v},\mathbf{v}) \le C\|\mathbf{v}\|_{-1}^2, \quad \forall\,\mathbf{v}\in L^2(\Omega)^2 \tag{4.10}$$

which can be easily proved by using the approximation property and (4.8).

Lemma 4.1. *For any* $(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p)\in(\tilde{\mathbf{L}}^2(\Omega)\cap H(\mathrm{div};\Omega)^2)\times H_0^1(\Omega)^2\times(L_0^2(\Omega)\cap H^1(\Omega))$, *there exists a constant* C, *independent of* h, μ *and* λ, *such that*

$$\frac{1}{C}\;|||(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p)|||^2 \;\le\; G_h(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p;\mathbf{0}) \tag{4.11}$$

and

$$G_h(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p;\mathbf{0}) \le C\left(\mu\|\underset{\sim}{\boldsymbol{\sigma}}\|^2 + \mu^2\|\mathbf{u}\|_1^2 + \lambda^2\|p\|^2 + \mu h^2\|\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}}\|^2 + \lambda^2h^2\|\nabla p\|^2\right). \tag{4.12}$$

If, in addition, $(\underset{\sim}{\boldsymbol{\sigma}},\mathbf{u},p)\in\boldsymbol{\mathcal{V}}^h = \underset{\sim}{\Phi}^h\times\mathcal{U}^h\times P^h$ *and if the spaces* $\underset{\sim}{\Phi}^h$ *and* P^h *satisfy inverse inequalities of the form*

$$\|\nabla\cdot\underset{\sim}{\boldsymbol{\sigma}}\| \le Ch^{-1}\|\underset{\sim}{\boldsymbol{\sigma}}\| \quad \textit{and} \quad \|\nabla p\| \le Ch^{-1}\|p\|, \tag{4.13}$$

respectively, then (4.12) can be replaced by

$$G_h(\underset{\sim}{\sigma}, \mathbf{u}, p; \mathbf{0}) \ \leq \ C \ |||(\underset{\sim}{\sigma}, \mathbf{u}, p)|||^2. \tag{4.14}$$

Proof. From (4.5) and (4.10), we have that

$$(T_h \mathbf{v}, \mathbf{v}) \leq C \left(h^2 \|\mathbf{v}\|^2 + (\tilde{B}_h \mathbf{v}, \mathbf{v})\right) \leq C \left(h^2 \|\mathbf{v}\|^2 + \|\mathbf{v}\|_{-1}^2\right), \quad \forall\, \mathbf{v} \in L^2(\Omega)^2.$$

The last inequality, the triangle inequality and Theorem 3.1 imply the inequality (4.12) and then the upper bound (4.14) is nothing but a combination of (4.12) and (4.13). We now prove the lower bound (4.11). By Theorem 3.1, it is enough to show that

$$\|\sqrt{2\mu}\nabla \cdot \underset{\sim}{\sigma} - \lambda \nabla p\|_{-1}^2 \leq C \left(T_h(\sqrt{2\mu}\nabla \cdot \underset{\sim}{\sigma} - \lambda \nabla p), \sqrt{2\mu}\nabla \cdot \underset{\sim}{\sigma} - \lambda \nabla p \right)$$

for any $\underset{\sim}{\sigma} \in (\tilde{\mathbf{L}}^2(\Omega) \cap H(\mathrm{div};\Omega)^2)$ and any $p \in H^1(\Omega)$. The last inequality can be proved by showing the following inequality. Using (4.9) and (4.10) we have

$$\|\mathbf{v}\|_{-1}^2 \leq 2\left(\|(I - Q_h)\mathbf{v}\|_{-1}^2 + \|Q_h \mathbf{v}\|_{-1}^2\right) \leq C \left(h^2 \|\mathbf{v}\|^2 + (\tilde{B}_h \mathbf{v}, \mathbf{v})\right) \leq C(T_h \mathbf{v}, \mathbf{v}),$$

for any $\mathbf{v} \in L^2(\Omega)^2$. This completes the proof of the lemma.

Theorem 4.2. *Assume that the spaces $\underset{\sim}{\Phi}^h$ and P^h satisfy the inverse inequalities (4.13). Let $(\underset{\sim}{\sigma}^h, \mathbf{u}^h, p^h)$ in $\boldsymbol{\mathcal{V}}^h$ be the unique minimizer of the problem (4.7). If the solution $(\underset{\sim}{\sigma}, \mathbf{u}, p)$ of the problem (2.4) satisfies $(\underset{\sim}{\sigma}, \mathbf{u}, p) \in \left(H^r(\Omega)^{2\times 2} \times H_0^{r+1}(\Omega)^2 \times H^r(\Omega)\right)$, then there exists a constant C, independent of h, μ and λ, such that*

$$|||(\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^h, \mathbf{u} - \mathbf{u}^h, p - p^h)||| \ \leq \ C\, h^r \ \left(\sqrt{\mu}\, \|\underset{\sim}{\sigma}\|_r + \mu\, \|\mathbf{u}\|_{r+1} + \lambda\, \|p\|_r \right). \tag{4.15}$$

Proof. It is easy to show that the error $(\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^h, \mathbf{u} - \mathbf{u}^h, p - p^h)$ is orthogonal to $\boldsymbol{\mathcal{V}}^h$ with respect to the inner product corresponding to the functional $G_h(\underset{\sim}{\sigma}, \mathbf{u}, p; \mathbf{0})$. Then the approximation properties (4.1-4.3), (4.11) and (4.14) in Lemma 4.1 yield the error estimate (4.15).

5. Implementations and Numerical Results

In this section we present the numerical experiments for the discrete H^{-1} least-squares methods (4.7). For the finite element spaces, we use the piecewise linear finite elements for the approximation of all unknowns. The use of a single approximating space for all variables simplifies the programming of least-squares finite element methods. The computational domain is taken to be the unit square. We use the uniform triangulation which has the grid interval h for each direction. For fixed $\mu = 1$, the numerical experiments were performed for $\lambda = 1, 10, 100, 1000$. For convenience, the unknowns are ordered as:

$$\mathbf{U} = (\sigma_1, \sigma_2, \sigma_3, u_1, u_2, p)^t \in \boldsymbol{\mathcal{V}}^h.$$

We will use B_h as the preconditioner for $\tilde{B}_h$ corresponding to one sweep of the multigrid V-cycle algorithm associated with (4.4) using the point Gauss-Seidel smoothing iteration. This multigrid preconditioner results from applying one step of the iterative procedure with zero starting iterate and uses one pre and post Gauss-Seidel iteration sweep.

Now, the problem (4.7) can be written as the matrix problem such that

$$\mathbf{A}_h \hat{\mathbf{U}}^h = \hat{\mathbf{F}}^h \tag{5.1}$$

where $\hat{\mathbf{U}}^h$ denotes the coefficient vector of $\mathbf{U}^h \in \boldsymbol{\mathcal{V}}^h$. Here, the matrix $\mathbf{A}_h$ itself is never assembled because of the operator T_h appeared in $\mathcal{A}_h$. Using the equivalent relation provided in (4.11) and (4.14) yields that the matrix $\mathbf{A}_h$ is spectrally equivalent to

$$\mathbf{R}_h := \mathrm{diag}[\,\mu h^2 I_\sigma,\, \mu^2 N_u,\, \lambda^2 h^2\,]$$

where I_σ and I_p denote the identity matrices associated with $\underset{\sim}{\Phi}^h$ and P^h, respectively, and N_u denotes the stiffness matrix associated with (4.4). We use the inverse of the block matrix $\mathbf{R}_h$ as a preconditioner for $\mathbf{A}_h$, in which the inverse of N_u itself is never assembled. The multigrid V-cycle preconditioner B_h will be used instead of computing N_u^{-1}.

To study multigrid V-cycle performance we use the multigrid V-cycle algorithm V(1,1) with the following relaxation sweep: for a given $\hat{\mathbf{V}}_0^h$,

$$\hat{\mathbf{V}}_{n+1}^h = \hat{\mathbf{V}}_n^h + \mathbf{R}_h^{-1}\,(\hat{\mathbf{F}}^h - \mathbf{A}_h\,\hat{\mathbf{V}}_n^h), \quad n \geq 0,$$

in which the grid size of the coarsest mesh is $h_0 = \frac{1}{2}$. The parameters α and β in $T_h = \alpha\ h^2 I + \beta\ B_h$ could be used to tune the iterative convergence rate. In all of the reported computations, we took $(\alpha, \beta) = (0.1, 0.8)$.

Define a discrete matrix norm $|||\hat{\mathbf{U}}^h|||_h$ of the coefficient $\hat{\mathbf{U}}^h = (\hat{\underset{\sim}{\boldsymbol{\sigma}}}^h, \hat{\mathbf{u}}^h, \hat{p}^h)$ of $\mathbf{U}^h = (\underset{\sim}{\boldsymbol{\sigma}}^h, \mathbf{u}^h, p^h) \in \boldsymbol{\mathcal{V}}^h$ by

$$|||\hat{\mathbf{U}}^h|||_h = \left(\mu\,\|\hat{\underset{\sim}{\boldsymbol{\sigma}}}^h\|_h^2 + \mu^2\,\|\hat{\mathbf{u}}^h\|_{1,h}^2 + \lambda^2\,\|\hat{p}^h\|_h^2\right)^{\frac{1}{2}}$$

where $\|\cdot\|_h$ and $\|\cdot\|_{1,h}$ denote the discrete L^2- and H^1-norms given by

$$\|\hat{\underset{\sim}{\boldsymbol{\sigma}}}^h\|_h = \sqrt{(\hat{\underset{\sim}{\boldsymbol{\sigma}}}^h)^t M_\sigma(\hat{\underset{\sim}{\boldsymbol{\sigma}}}^h)},\ \|\hat{\mathbf{u}}^h\|_{1,h} = \sqrt{(\hat{\mathbf{u}}^h)^t N_u(\hat{\mathbf{u}}^h)} \text{ and } \|\hat{p}^h\|_h = \sqrt{(\hat{p}^h)^t M_p(\hat{p}^h)}$$

with the mass matrices M_σ and M_p and the stiffness matrix N_u corresponding to the spaces $\underset{\sim}{\Phi}^h$, P^h and $\mathcal{U}^h$, respectively.

Example 1. To study multigrid V-cycle performance we begin by studying the convergence factors for the trivial problem which has zero exact solution, i.e., we take $\mathbf{f} = 0$ in problem (2.4). With initial guess one, i.e., $\hat{\mathbf{U}}_0^h = 1$, we define the convergence factors ρ_k^h as ratios of discrete matrix norm of successive iterations such that

$$\rho_k^h := \frac{|||\hat{\mathbf{U}}_{k+1}^h|||_h}{|||\hat{\mathbf{U}}_k^h|||_h}, \qquad \forall\, k \geq 0.$$

In Table 1, we present the convergence factors for $\lambda = 1, 10, 100, 1000$ measured after 20 V(1,1)-cycles, in which the linear finite elements was used. The table shows that the convergence factors appear to be independent of h and λ.

Example 2. In order to measure discretization errors by discrete H^{-1}-norm least-squares method we chose a problem with a known nonzero solution to the problem (2.4):

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} x(1-x)y^2(1-y)^2 \sin \pi x \\ x^2(1-x)^2 y^2(1-y)^2 \cos \pi y \end{pmatrix}.$$

Then, from this displacement solution $\mathbf{u}$ we obtained the other solutions $\underset{\sim}{\sigma}$ and p and the data function $\mathbf{f}$ in (2.4), Table 2 shows the total errors $|||(\mathbf{e}_\sigma, \mathbf{e}_u, \mathbf{e}_p)|||_h$ and their convergence rates where $\mathbf{e}_\sigma, \mathbf{e}_u$ and $\mathbf{e}_p$ denote the differences between the exact solutions and their approximations. The table shows that the error bounds are like $O(h^{2-\epsilon})$ with $0 < \epsilon < \frac{1}{2}$ for $\lambda = 1, 10$ and almost $O(h^2)$ for $\lambda = 100, 1000$, but the theoretically predicted error bounds are $O(h)$. Therefore we may have obtained superconvergence for the use of linear elements.

In order to compare our numerical results with best approximation, in the remaining tables we show the L^2 and H^1 discretization errors and their rates along with three components for $\lambda = 1, 10, 100, 1000$. The rates of the best approximation out of the linear finite element spaces are $O(h^2)$ in L^2 and $O(h)$ in H^1. In Tables 4 through 7, we observe that the computed L^2 convergence rates of $\mathbf{e}_u$ for $\lambda = 1, 10, 100, 1000$ and $\mathbf{e}_\sigma$ and $\mathbf{e}_p$ for $\lambda = 100, 1000$ are almost in agreement with the best rates $O(h^2)$. In those tables, we can see a common phenomenon that the larger λ is, the better the convergence rates of each variable in both L^2 and H^1 are. Tables 8 through 11 provide the H^1 discretization errors and their rates. We note that the theoretical error bounds of $\mathbf{e}_u$ in H^1 is $O(h)$ and the best error bounds for all variables in H^1 is $O(h)$. The tables show that the rates of convergence for the displacement errors $\mathbf{e}_u$ are like $O(h^{2-\epsilon})$.

Table 1. Convergence factors after 20 MGV(1,1)-cycles

	$\lambda = 1$	$\lambda = 10$	$\lambda = 100$	$\lambda = 1000$
$h = \frac{1}{8}$	0.8868	0.8599	0.8677	0.8590
$h = \frac{1}{16}$	0.8794	0.8426	0.8734	0.8710
$h = \frac{1}{32}$	0.8694	0.8417	0.8768	0.8756
$h = \frac{1}{64}$	0.8624	0.8412	0.8759	0.8767
$h = \frac{1}{128}$	0.8579	0.8406	0.8704	0.8767
$h = \frac{1}{256}$	0.8552	0.8401	0.8618	0.8764

Table 2. Total errors $|||(\mathbf{e}_\sigma^h, \mathbf{e}_u^h, \mathbf{e}_p^h)|||_h$ and convergence rates

	$\lambda = 1$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	6.56e-3		9.03e-3		7.61e-2		7.31e-1	
$h = \frac{1}{16}$	2.54e-3	1.37	3.20e-3	1.50	8.59e-3	3.15	7.90e-2	3.21
$h = \frac{1}{32}$	7.99e-4	1.67	9.14e-4	1.81	1.63e-3	2.40	1.14e-2	2.79
$h = \frac{1}{64}$	2.36e-4	1.76	2.66e-4	1.78	4.34e-4	1.91	2.64e-3	2.11
$h = \frac{1}{128}$	7.14e-5	1.73	7.76e-5	1.78	1.14e-4	1.92	6.65e-4	1.99
$h = \frac{1}{256}$	2.26e-5	1.66	2.36e-5	1.72	3.07e-5	1.90	1.72e-4	1.96

Table 3. L^2-norm errors and convergence rates when $\lambda = 1$

	$\mathbf{e}_\sigma^h$		$\mathbf{e}_\mathbf{u}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	3.048e-3		8.786e-4		1.126e-3	
$h = \frac{1}{16}$	1.190e-3	1.3564	2.990e-4	1.5547	5.196e-4	1.1157
$h = \frac{1}{32}$	3.759e-4	1.6629	8.487e-5	1.8171	2.268e-4	1.1960
$h = \frac{1}{64}$	1.093e-4	1.7813	2.223e-5	1.9324	8.448e-5	1.4247
$h = \frac{1}{128}$	3.216e-5	1.7655	5.728e-6	1.9567	3.004e-5	1.4917
$h = \frac{1}{256}$	9.894e-6	1.7009	1.493e-6	1.9393	1.056e-5	1.5083

Table 4. L^2-norm errors and convergence rates when $\lambda = 10$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	3.201e-3		6.984e-4		6.206e-4	
$h = \frac{1}{16}$	1.372e-3	1.2217	3.022e-4	1.2082	8.644e-5	2.8439
$h = \frac{1}{32}$	4.479e-4	1.6155	8.274e-5	1.8691	3.253e-5	1.4099
$h = \frac{1}{64}$	1.260e-4	1.8294	2.178e-5	1.9254	1.058e-5	1.6204
$h = \frac{1}{128}$	3.499e-5	1.8488	5.526e-6	1.9790	3.410e-6	1.6335
$h = \frac{1}{256}$	1.006e-5	1.7984	1.387e-6	1.9937	1.130e-6	1.5934

Table 5. L^2-norm errors and convergence rates when $\lambda = 100$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	3.500e-2		1.238e-3		6.592e-4	
$h = \frac{1}{16}$	3.688e-3	3.2467	2.815e-4	2.1369	6.937e-5	3.2483
$h = \frac{1}{32}$	5.609e-4	2.7170	8.408e-5	1.7432	1.292e-5	2.4247
$h = \frac{1}{64}$	1.498e-4	1.9040	2.276e-5	1.8850	3.380e-6	1.9345
$h = \frac{1}{128}$	4.034e-5	1.8936	5.939e-6	1.9385	8.500e-7	1.9915
$h = \frac{1}{256}$	1.079e-5	1.9018	1.521e-6	1.9651	2.100e-7	2.0171

Table 6. L^2-norm errors and convergence rates when $\lambda = 1000$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	3.349e-1		8.150e-3		6.411e-4	
$h = \frac{1}{16}$	3.323e-2	3.3334	9.521e-4	3.0975	6.973e-5	3.2007
$h = \frac{1}{32}$	2.442e-3	3.7663	1.242e-4	2.9377	1.095e-5	2.6708
$h = \frac{1}{64}$	2.217e-4	3.4612	2.818e-5	2.1405	2.610e-6	2.0688
$h = \frac{1}{128}$	4.469e-5	2.3106	7.122e-6	1.9845	6.600e-7	1.9835
$h = \frac{1}{256}$	1.198e-5	1.8988	1.786e-6	1.9950	1.700e-7	1.9569

Table 7. H^1-norm errors and convergence rates when $\lambda = 1$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	2.945e-2		5.634e-3		1.356e-2	
$h = \frac{1}{16}$	1.672e-2	0.8163	2.157e-3	1.3848	1.532e-2	-0.176
$h = \frac{1}{32}$	1.137e-2	0.5566	6.621e-4	1.7045	1.433e-2	0.0964
$h = \frac{1}{64}$	8.042e-3	0.4996	1.902e-4	1.7992	1.095e-2	0.3881
$h = \frac{1}{128}$	5.676e-3	0.5027	5.589e-5	1.7671	7.901e-3	0.4708
$h = \frac{1}{256}$	3.992e-3	0.5076	1.727e-5	1.6944	5.598e-3	0.4971

Table 8. H^1-norm errors and convergence rates when $\lambda = 10$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	Error	Rate	Error	Rate	Error	Rate
$h = \frac{1}{8}$	4.135e-2		5.680e-3		6.834e-3	
$h = \frac{1}{16}$	1.809e-2	1.1927	2.744e-3	1.0493	1.533e-3	2.1564
$h = \frac{1}{32}$	1.116e-2	0.6963	7.220e-4	1.9266	1.476e-3	0.0547
$h = \frac{1}{64}$	7.268e-3	0.6192	2.079e-4	1.7960	1.065e-3	0.4708
$h = \frac{1}{128}$	4.998e-3	0.5403	6.004e-5	1.7920	7.566e-4	0.4932
$h = \frac{1}{256}$	3.500e-3	0.5139	1.811e-5	1.7287	5.355e-4	0.4986

Table 9. H^1-norm errors and convergence rates when $\lambda = 100$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	ERROR	RATE	ERROR	RATE	ERROR	RATE
$h = \frac{1}{8}$	4.789e-1		1.497e-2		9.121e-3	
$h = \frac{1}{16}$	6.675e-2	2.8429	3.473e-3	2.1076	1.551e-3	2.5560
$h = \frac{1}{32}$	1.438e-2	2.2144	8.240e-4	2.0755	4.085e-4	1.9248
$h = \frac{1}{64}$	6.307e-3	1.1891	2.254e-4	1.8702	1.487e-4	1.4579
$h = \frac{1}{128}$	3.608e-3	0.8059	6.479e-5	1.7987	6.762e-5	1.1369
$h = \frac{1}{256}$	2.343e-3	0.6225	1.952e-5	1.7305	3.815e-5	0.8258

Table 10. H^1-norm errors and convergence rates when $\lambda = 1000$

	$\mathbf{e}_{\boldsymbol{\sigma}}^h$		$\mathbf{e}_{\mathbf{u}}^h$		$\mathbf{e}_p^h$	
	ERROR	RATE	ERROR	RATE	ERROR	RATE
$h = \frac{1}{8}$	4.640e+0		1.020e-1		8.911e-3	
$h = \frac{1}{16}$	6.230e-1	2.8970	1.643e-2	2.6342	1.351e-3	2.7216
$h = \frac{1}{32}$	7.878e-2	2.9833	2.025e-3	3.0203	2.462e-4	2.4561
$h = \frac{1}{64}$	1.217e-2	2.6945	3.053e-4	2.7299	5.583e-5	2.1407
$h = \frac{1}{128}$	4.770e-3	1.3513	7.237e-5	2.0768	1.758e-5	1.6671
$h = \frac{1}{256}$	3.171e-3	0.5888	2.041e-5	1.8257	7.410e-6	1.2464

References

[1] I. Babuška and B. Szabo, On the rates of convergence of the finite element method, *Internat. J. Numer. Meth. Engng.,* **18** (1982), 323–341.

[2] M. A. Behr, L. P. Franca and T. E. Tezduyar, Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows, *Comput. Methods Appl. Mech. Engrg.,* **104** (1993), 31–48.

[3] P. B. Bochev and M. D. Gunzburger, Least-squares methods for the velocity-pressure-stress formulation of the Stokes equations, *Comput. Methods in Appl. Mech. Engrg.,* **126** (1995), 267–287.

[4] P. B. Bochev and M. D. Gunzburger, Analysis of least-squares finite element methods for the Stokes equations, *Comput. Fluids,* **22** (1993), 549–563.

[5] P. B. Bochev and M. D. Gunzburger, Accuracy of least-squares methods for the Navier-Stokes equations, *Math. Comp.,* **63** (1994), 479–506.

[6] J. Bramble, R. Lazarov and J. Pasciak, A least squares approach based on a discrete minus one inner product for first order systems, Technical report 94-32, Mathematical Science Institute, Cornell University, Ithaca, NY, 1994.

[7] J. Bramble and J. Pasciak, Least squares method for Stokes equations based on a discrete minus one inner product, *J. Comp. Appl. Math.,* **74** (1996), 155–173.

[8] S. C. Brenner and L. Y. Sung, Linear finite element method for planar linear elasticity, *Math. Comp.,* **59** (1992), 321–338.

[9] Z. Cai, C.-O. Lee, T. A. Manteuffel and S. F. McCormick, First-order system least-squares for linear elasticity: Numerical results, *SIAM J. Sci. Comput.,* **21** (2000), 1706–1727.

[10] Z. Cai, C.-O. Lee, T. A. Manteuffel and S. F. McCormick, First-order system least-squares for linear elasticity: Further results, *SIAM J. Sci. Comput.,* **21** (2000), 1728–1739.

[11] Z. Cai, T. Manteuffel and S. McCormick, First-order system least squares for the Stokes equations, with application to linear elasticity, *SIAM J. Numer. Anal.,* **34** (1997), 1727–1741.

[12] Z. Cai, T. Manteuffel, S. McCormick and S. V. Parter, First-order system least squares(FOSLS) for planar linear elasticity: pure traction problem, *SIAM J. Numer. Anal.,* **35** (1998), 320–225.

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

[14] P. G. Ciarlet, The finite element method for elliptic problems, North-Holland, New York, 1978.

[15] L. P. Franca and R. Stenberg, Error analysis of some Galerkin least-squares methods for the elasticity equations, *SIAM J. Numer. Anal.,* **28** (1991), 1680–1697.

[16] V. Girault and P. A. Raviart, Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms, Springer-Verlag, New York, 1986.

[17] C. Johnson, A mixed finite element method for the Navier-Stokes equations, *RAIRO* M^2AN, **12** (1978), 335–348.

[18] S. D. Kim, T. A. Manteuffel and S. F. McCormick, First-order system least squares (FOSLS) for spatial linear elasticity: pure traction, *SIAM J. Numer. Anal.,* **38** (2000), 1454–1482.

[19] J. Nečas, Equations aux Dérivées Partielles, Presses de l'Université de Montréal, 1965.

[20] M. Vogelius, An analysis of the p-version of the finite element method for nearly incompressible materials: uniformly valid, optimal error estimates, *Numer. Math.,* **41** (1983), 39–53.

[21] S.-Y. Yang and J.-L. Liu, Least-squares finite element methods for the elasticity problem, *Jour. Comp. Appl. Math.,* **87** (1997), 39–60.