$H^{-1}$ least-squares method for the velocity–pressure–stress formulation of Stokes equations

Sang Dong Kim $^{a,*,1}$, Byeong Chun Shin $^{b,2}$

$^a$ Department of Mathematics, Kyungpook National University, Taegu, Republic of Korea 702-701
$^b$ Department of Mathematics, Seoul National University, Seoul, Republic of Korea 151-747

Abstract

This paper studies the discrete $H^{-1}$-norm least-squares method for the incompressible Stokes equations based on the velocity–pressure–stress formulation by the least-squares functional defined as the sum of $L^2$-norms and $H^{-1}$-norm of the residual equations. Some computational experiments by multigrid method and preconditioning conjugate gradient method (PCGM) on this method are shown by taking efficient $\alpha$ and $\beta$ in the discrete solution operator $T_h = \alpha h^2 I + \beta B_h$ corresponding to the minus one norm. We also propose a new method and compare it with PCGM and multigrid method through the analysis of numerical experiments depending on the choice of $\beta$.

Keywords: Least-squares method; Stokes equations; Multigrid method; Preconditioning conjugate gradient method

1. Introduction

In recent years there have been lots of developments in the use of least-squares methods for numerical approximations of the incompressible Stokes and Navier–Stokes equations based on velocity–flux–pressure, velocity–vorticity–pressure and velocity–pressure–stress formulations [2–8,12,13]. In recent works by Bochev and Gunzburger, the ADN approach was extended to the vorticity formulation [4] and to the stress formulation [6] of the Stokes equations. The least-squares functional they used is defined to be the sum of square norms of the residuals, where the norms are determined by the indices assigned to each equation by the ADN theory [1]. They replace the $H^1$-norms by mesh-dependent $L^2$-norms, $h^{-1} \| \cdot \|_0$. Cai et al. analyzed the $H^{-1}$-norm least-squares functional for Stokes equations based on the flux formulation [12] and vorticity formulation [13], in which its least-squares functional is similar to...
that in [6] with $q = -1$ and the $H^{-1}$-norm of the functional can be replaced by the discrete $H^{-1}$-norm for the feasible computation following the idea introduced by Bramble et al. [9].

In this paper, we consider the discrete $H^{-1}$-norm least-squares method for the Stokes equations based on the velocity–pressure–stress formulation. The analysis of the discrete $H^{-1}$-norm least-squares methods is similar to that in [13]. Hence, we focus on the development of an efficient preconditioner for the resulting linear algebraic system and the computational analysis of this method by analyzing an efficient role of $\alpha$ and $\beta$ in the discrete operator $T_h = \alpha h^2 I + \beta B_h$ used in the definition of the discrete minus one norm and discuss performances by multigrid method where $B_h$ is the multigrid V-cycle preconditioner. We analyze some numerical experiments including the convergence factors by PCGM and multigrid method with the several choices of $\beta$. Through the analysis of numerical results, we also suggest a new perturbed solution method depending on $\beta$, which is compared with PCGM and multigrid method in terms of iteration numbers. Several computational convergence rates are provided for all unknowns when the continuous piecewise linear finite elements for the approximations of all variables are used. The theoretically predicted discretization error bounds are $O(h)$ in $L^2$ for the stress and pressure and $O(h)$ in $H^1$ for the velocity. The resulting convergence rates for the velocity asymptotically approach to the best approximation rates $O(h^2)$ in $L^2$ and $O(h^{3/2+\delta})$ in $H^1$ and the rates for the stress and pressure are like $O(h^{3/2+\delta})$ in $L^2$ and $O(h^{1/2})$ in $H^1$, from which it appears that we obtain an optimal convergence in the $L^2$-norm and super-convergence in the $H^1$-norm for the velocity but suboptimal convergence rates for the stress and pressure. In order to obtain the optimal convergence rates for all unknowns, one may use polynomials of one degree lower for the stress and pressure than that for the velocity. However, the use of a single approximating space for all variables simplifies the programming of least-squares finite element methods.

This paper is organized as follows. In Section 2, we give an equivalent first-order system of linear equations to the Stokes equations and define the $H^{-1}$-norm least-squares method for the stress formulation. The ellipticity and continuity for the discrete $H^{-1}$-norm least-squares functional are established and its finite element approximation with an error estimate is given in Section 3. In Section 4, we study a preconditioner for the resulting linear system and analyze several numerical experiments performed by PCGM and multigrid method.

2. First-order system $H^{-1}$ least-squares methods

Let $\Omega$ be a bounded, open, connected domain in $\mathbb{R}^d$ ($d = 2$ or 3) with Lipschitz boundary $\partial \Omega$. Consider the steady-state incompressible Stokes equation of the form

\begin{align}
-\nu \Delta u + \nabla p &= f \quad \text{in } \Omega, \\
\nabla \cdot u &= 0 \quad \text{in } \Omega, \\
\qquad u = u_0 \quad \text{on } \partial \Omega, \tag{2.1}
\end{align}

where $u$ denotes the velocity, $p$ the pressure, $\nu > 0$ the reciprocal of the Reynolds number and $f$ is the body force. We assume that the pressure $p$ satisfies a zero mean constraint

\begin{align}
\int_{\Omega} p \, dx &= 0. \tag{2.2}
\end{align}
For convenience, let the boldface denote the vector valued function and the undertilde boldface the matrix-valued function, i.e., the tensor. The colon notation: denotes the inner product on $\mathbb{R}^{d \times d}$ and for any tensors $\tau = (\tau_{ij})$ and $\delta = (\delta_{ij})$ in $L^2(\Omega)^{d \times d}$, the $L^2(\Omega)^{d \times d}$ inner product is defined by

$$(\tau, \delta) = \int_\Omega \tau : \delta \, dx.$$ 

The divergence for a tensor $\tau$ is defined as

$$\nabla \cdot \tau = \left( \sum_{j=1}^d \frac{\partial \tau_{1j}}{\partial x_j}, \sum_{j=1}^d \frac{\partial \tau_{2j}}{\partial x_j}, \ldots, \sum_{j=1}^d \frac{\partial \tau_{dj}}{\partial x_j} \right)^T,$$

where $^T$ denotes transpose. Let $\varepsilon(u)$ denote the symmetric part of the velocity gradient, i.e., the deformation tensor

$$\varepsilon(u) := \frac{1}{2} \left( \nabla u + \nabla u^T \right) = (\varepsilon_{i,j}), \quad \varepsilon_{i,j} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, \ldots, d.$$

Defining the stress tensor $\sigma := \sqrt{2\nu} \varepsilon(u)$ scaled by $\sqrt{\nu}$, we have the following generalized velocity–pressure–stress system (see [6])

$$\sigma - \sqrt{2\nu} \varepsilon(u) = F_1 \quad \text{in} \ \Omega,$$

$$\sqrt{2\nu} \nabla \cdot \sigma - \nabla p = -f_2 \quad \text{in} \ \Omega,$$

$$\nabla \cdot u = f_3 \quad \text{in} \ \Omega,$$

$$u = u_0 \quad \text{on} \ \partial \Omega,$$

where the function $f_3$ satisfies the following solvability constraint:

$$\int_\Omega f_3 \, dx = \int_{\partial \Omega} u_0 \cdot n \, ds. \quad (2.4)$$

In two dimensions, (2.3) is a system of six unknowns and six equations and in three dimensions the number of the unknowns and the equations increase to ten. If the tensor $F_1$ and the function $f_3$ are identically zero, the Stokes problem (2.1) is equivalent to the generalized system (2.3). For simplicity, without loss of generality, we assume that $u_0 = 0$. Then, (2.4) implies that $f_3$ has zero mean.

Let $H^s(\Omega)$ be the Sobolev spaces with the standard associated inner products $(\cdot, \cdot)_s$ and their respective norms $\| \cdot \|_s$. For $s = 0$, $H^1(\Omega)$ coincides with $L^2(\Omega)$. In this case the norm and inner product will be denoted by $\| \cdot \|$ and $(\cdot, \cdot)$, respectively. $H^s_0(\Omega)$ denotes the closure of $D(\Omega)$, the linear space of infinitely differentiable functions with compact supports on $\Omega$, with respect to the norm $\| \cdot \|_s$. Denote by $L^2_0(\Omega)$ the subspace of square integrable functions with zero mean:

$$L^2_0(\Omega) := \left\{ p \in L^2(\Omega) : \int_\Omega p \, dx = 0 \right\}.$$
For positive values of $s$ the space $H^{-s}(\Omega)$ is defined as the dual space of $H^s_0(\Omega)$ equipped with the norm

$$
\|\phi\|_{-s} := \sup_{0 \neq v \in H^s_0(\Omega)} \frac{(\phi, v)}{\|v\|_s},
$$

where $(\cdot, \cdot)$ is the duality pairing between $H^s(\Omega)$ and $H^s_0(\Omega)$ when there is no risk of confusion. Define the product spaces $H^s_i(\Omega)^d = \prod_{i=1}^d H^s_0(\Omega)$ and $H^{-s}(\Omega)^d = \prod_{i=1}^d H^{-s}(\Omega)$ with standard product norms. Let $H(\text{div}; \Omega) = \{v \in L^2(\Omega)^d: \nabla \cdot v \in L^2(\Omega)\}$ which is Hilbert space under the norm $\|v\|_{H(\text{div}; \Omega)} := (\|v\|^2 + \|\nabla \cdot v\|^2)^{1/2}$.

Let $T : H^{-1}(\Omega)^d \rightarrow H^1_0(\Omega)^d$ denote the solution operator such that for any $f \in H^{-1}(\Omega)^d$, $T f = w$ is the solution to the following boundary value problem:

\begin{align*}
-\Delta w + w &= f & \text{in } \Omega, \\
w &= 0 & \text{on } \partial \Omega.
\end{align*}

(2.5)

From the definition of $T$, we can easily show that

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

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

Taking account of (2.6), we define a least-squares functional as the weighted sum of the $L^2$- and $H^{-1}$-norms of the residual equations of system (2.3):

$$
G(\sigma, u; f) = (T(\sqrt{2\nu} \nabla \cdot \sigma - \nabla p + f), \sqrt{2\nu} \nabla \cdot \sigma - \nabla p + f)
$$

$$
+ 2\nu \|\sigma - \sqrt{2\nu} \nu u - F_1\|^2 + 4\nu^2 \|\nabla \cdot u - F_3\|^2.
$$

(2.7)

Let

$$
\mathcal{V} := L^2(\Omega) \times H^1_0(\Omega)^d \times L^2(\Omega),
$$

where $L^2(\Omega)$ is the space of $d \times d$ symmetric matrix functions whose elements are square-integrable. Then, the first-order system least-squares variational problem for the Stokes equations is to minimize the quadratic functional $G(\sigma, u; p; f)$ over $\mathcal{V}$: find $(\sigma, u, p) \in \mathcal{V}$ such that

$$
G(\sigma, u, p; f) = \inf_{(\tau, v, q) \in \mathcal{V}} G(\tau, v, q; f).
$$

(2.8)

By the arguments similar to Theorem 2.1 in [13] we can easily establish ellipticity of the functional $G(\sigma, u, p; 0)$. As a result, our functional is robust in a sense of the viscosity $\nu$.

**Theorem 1.** For any $(\sigma, u, p) \in \mathcal{V}$, there exist constants $c$ and $C$, independent of $\nu$ such that

$$
c(2\nu \|\sigma\|^2 + 4\nu^2 \|u\|_1^2 + \|p\|^2) \leq G(\sigma, u, p; 0)
$$

(2.9)

and

$$
G(\sigma, u, p; 0) \leq C(2\nu \|\sigma\|^2 + 4\nu^2 \|u\|_1^2 + \|p\|^2).
$$

(2.10)
A similar functional without the weights of the parameter \( v \) for the Stokes equations was considered by Bochev and Gunzburger [6] and they applied a priori error estimate by ADN theory to establish ellipticity of the functional

\[
\| \sigma \|_{q+1} + \| p \|_{q+1} + \| u \|_{q+2} \leq C (\| \sqrt{2} v \nabla \cdot \sigma - \nabla p \|_q + \| \nabla \cdot u \|_{q+1} + \| \sigma - \sqrt{2} \nabla \cdot \mathbf{g}(u) \|_{q+1}).
\]

Our \( H^{-1} \)-norm least-squares functional is similar to the right hand side of the above inequality when \( q = -1 \).

3. Finite element approximation

In this section we present a discrete \( H^{-1} \)-norm least-squares finite element approximations, the well-posedness of the discrete problem and then establish optimal order error estimate for each variable. For the finite element approximation, we assume that the domain \( \Omega \) is a polygonal for \( d = 2 \) or a polyhedron for \( d = 3 \) and that \( T_h \) is a partition of \( \Omega \) into finite elements \( \Omega = \bigcup_{K \in T_h} K \) with \( h = \max \{ \text{diam}(K) : K \in T_h \} \). Assume that the triangulation \( T_h \) is regular and satisfies the inverse assumption (see [14]). Let \( \mathbf{V}^{h} := \mathbf{F}^{h} \times \mathcal{U}^{h} \times \mathcal{P}^{h} \) be a finite element subspace of \( \mathbf{V} \) with the following properties: there exist a constant \( C \) and an integer \( s \) such that for all \( (\sigma, u, p) \in (L^2(\Omega) \cap H^r(\Omega)^d) \times (H_0^1(\Omega)^d \cap H^{r+1}(\Omega)) \times (L_0^2(\Omega) \cap H^r(\Omega)) \), \( 1 \leq r \leq s \), there exists a triplet \( (\sigma^l, u^l, p^l) \in \mathbf{V}^{h} \) such that

\[
\inf_{\sigma^l \in \mathbf{F}^{h}} (\| \sigma - \sigma^l \| + h \| \sigma - \sigma^l \|_1) \leq Ch^r \| \sigma \|_r,
\]

\[
\inf_{u^l \in \mathcal{U}^{h}} (\| u - u^l \| + h \| u - u^l \|_1) \leq Ch^{r+1} \| u \|_{r+1}
\]

and

\[
\inf_{p^l \in \mathcal{P}^{h}} (\| p - p^l \| + h \| p - p^l \|_1) \leq Ch^r \| p \|_r.
\]

Note that typical finite element spaces consisting of continuous piecewise polynomials with respect to quasi-uniform triangulations satisfy (3.1)–(3.3) (see [14]).

Let \( (\sigma, u, p) \) be the solution of (2.3) which obviously minimizes the functional (2.7) and let \( (\sigma^h, u^h, p^h) \) minimize (2.7) over \( \mathbf{V}^{h} \). It follows from Theorem 1 that the bounds of the resulting errors \( e_\sigma = \sigma - \sigma^h, e_u = u - u^h \) and \( e_p = p - p^h \) are given by

\[
\sqrt{2} v e_\sigma + 2 v e_u + 2 v e_p \leq \inf_{(\tau, v, q) \in \mathbf{V}^{h}} \left\{ \sqrt{2} v \| \tau \| + 2 v \| u - v \|_1 + \| p - q \| \right\}.
\]

Although minimization with respect to the functional \( G(\cdot, \cdot, \cdot, \cdot, \cdot) \) appears attractive from the point of view of stability and accuracy, its computation is not feasible. This is because the evaluation of the operator \( T \) defining the inner product in \( H^{-1} \) involves the solution of the boundary value problem (2.5). Following the ideas suggested in [9] we will replace the operator \( T \) by an equivalent and computable operator \( T_h \). Let \( \tilde{B}_h : H^{-1}(\Omega)^d \to \mathcal{U}^{h} \) be the discrete solution operator \( \tilde{w} = \tilde{B}_h \psi \in \mathcal{U}^{h} \) for the Dirichlet problem (2.5) defined by

\[
(\nabla \tilde{w}, \nabla \tilde{v}) + (\tilde{w}, \tilde{v}) = (\psi, \tilde{v}), \quad \tilde{v} \in \mathcal{U}^{h}.
\]
Assume that there is a preconditioner \( B_h : H^{-1}(\Omega)^d \to \mathcal{U}_h \) for \( \tilde{B}_h \) that is symmetric positive definite operator with respect to the \( L^2(\Omega)^d \)-inner product and spectrally equivalent to \( B_h \), i.e., there exist positive constants \( c \) and \( C \) not depending on the mesh size \( h \) such that
\[
c(B_h v, v) \leq (\tilde{B}_h v, v) \leq C(B_h v, v), \quad v \in \mathcal{U}_h.
\]
(3.6)

For fixed positive constants \( \alpha \) and \( \beta \), define
\[
T_h = \alpha h^2 I + \beta B_h,
\]
(3.7)
where \( I \) denotes the identity operator on \( \mathcal{U}_h \). The positive parameters \( \alpha \) and \( \beta \) could be used to tune the iterative convergence rate. Now, let us define the discrete counterpart of the least-squares functional \( G(\cdot, \cdot) \) such that
\[
G_h(\sigma, u, p; f) = (T_h(\sqrt{2v} \nabla \cdot g - \nabla p + f_2), \sqrt{2v} \nabla \cdot g - \nabla p + f_2) + 2v\| \sigma - \sqrt{2v} g(u) - f_1 \|^2 + 4v^2 \| \nabla \cdot u - f_3 \|^2.
\]
(3.8)
The finite element approximation to (2.8) is to find \((\sigma^h, u^h, p^h) \in \mathcal{V}_h\) such that
\[
G_h(\sigma^h, u^h, p^h; f) = \inf_{(\sigma, u, p) \in \mathcal{V}_h} G_h(\sigma, u, p; f).
\]
(3.9)
Using the similar arguments in [9,13], we can easily derive the following lemma:

**Lemma 2.** For any \((\sigma, u, p) \in (L^2(\Omega) \cap H(\text{div}; \Omega)^d) \times H^1_0(\Omega) \times (L^2(\Omega) \cap H^1(\Omega))\), there exist \( c \) and \( C \), independent of \( h \) and \( v \), such that
\[
c(2v\| \sigma \|^2 + 4v^2\| u \|^2 + \| p \|^2) \leq G_h(\sigma, u, p; 0)
\]
(3.10)
and
\[
G_h(\sigma, u, p; 0) \leq C(2v\| \sigma \|^2 + 4v^2\| u \|^2 + \| p \|^2 + 2v^2\| \nabla \cdot g \|^2 + h^2\| \nabla p \|^2).
\]
(3.11)
If, in addition, \((\sigma, u, p) \in \mathcal{V}_h = \Phi^h \times \mathcal{U}_h \times P^h\) and if the spaces \( \Phi^h \) and \( P^h \) satisfy inverse inequalities of the form
\[
\| \nabla \cdot g \| \leq Ch^{-1}\| g \| \quad \text{and} \quad \| \nabla p \| \leq Ch^{-1}\| p \|,
\]
(3.12)
respectively, then (3.11) can be replaced by
\[
G_h(\sigma, u, p; 0) \leq (2v\| \sigma \|^2 + 4v^2\| u \|^2 + \| p \|^2).
\]
(3.13)

**Theorem 3.** Let \((\sigma, u, p) \in (H^r(\Omega)^{d_x d} \times H^r_0(\Omega)^d \times H^r(\Omega)) \cap \mathcal{V}\) be the solution of the problem (2.3) and let \((\sigma^h, u^h, p^h) \in \mathcal{V}_h\) be the solution of the problem (3.9). Then
\[
\sqrt{2v}\| \sigma - \sigma^h \| + 2v\| u - u^h \| + \| p - p^h \| \leq Ch^{r}(\sqrt{2v}\| \sigma \| + 2v\| u \| + \| p \|).
\]
(3.14)
Proof. It is easy to show that the error $(\sigma - \sigma^h, u - u^h, p - p^h)$ is orthogonal to $V^h$ with respect to the inner product corresponding to the functional $G_h(\sigma, u; p; 0)$. Then the approximation properties (3.1)–(3.3) and Lemma 2 yield the inequality (3.14).

If we use continuous piecewise quadratic polynomials for the approximation of the velocity and continuous piecewise linear polynomials for the approximations of the pressure and stress, then we get the following error estimate

$$\sqrt{2v}\|e_\sigma\| + 2\nu\|e_u\|_1 + \|e_p\| \leq C h^2 \{ \sqrt{2v}\|\sigma\|_2 + 2\nu\|u\|_3 + \|p\|_2 \},$$

which is optimal with respect to the finite element functions used. Note that the use of continuous piecewise linear polynomials for all unknowns yields the error estimate $O(h)$, but it does not give optimal error estimates for the pressure $p$ and stress $\sigma$. However, the use of a single approximating space for all variables simplifies programming of least-squares finite element methods.

4. Numerical experiments

In this section we present two dimensional $(d = 2)$ numerical experiments with the discrete $H^{-1}$ least-squares methods. We first study a preconditioner for the resulting algebraic system and then performances for preconditioning conjugate gradient method (PCGM) and multigrid V-cycle method (MGV) by condition numbers and convergence factors associated with the chosen parameter $\beta$. Through the analysis of the role of $\beta$ in the numerical results, we suggest a more efficient solution method. Finally, the finite element discretization accuracy is analyzed.

We take for our domain the unit square $\Omega = \{0 \leq x \leq 1, \ 0 \leq y \leq 1\}$ and consider the generalized velocity–pressure–stress Stokes equations (2.3). We choose the same exact solution discussed in [6]:

$$\begin{align*}
\sigma_1 &= \sigma_2 = \sigma_3 = \sin(\pi x) \exp(\pi y), \\
u_1 &= u_2 = \sin(\pi x) \sin(\pi y), \\
p &= \cos(\pi x) \exp(\pi y).
\end{align*}$$

Substituting this solution into the equations in (2.3) we define the data functions $F_1, f_2$ and $f_3$. The exact solution is smooth and satisfies the homogeneous velocity boundary condition. The computational domain is triangularized uniformly with the grid interval $h$ ranging from $2^{-2}$ to $2^{-8}$ for each direction. We use the single approximating space of continuous piecewise linear polynomials for the approximations of all unknowns. To define the discrete solution operator $T_h = \alpha h^2 I + \beta B_h$, we used $B_h$ corresponding to one sweep of the multigrid V(1, 1)-cycle algorithm associated with (3.5) using the point Gauss–Seidel smoothing iteration (see [9]). In this section, we use the following discrete $L^2$- and $H^1$-norm to measure the discretization errors,

$$\|V\|_h = \sqrt{V^T M_h V} \quad \text{and} \quad \|V\|_{1,h} = \sqrt{V^T S_h V},$$

where $M_h$ and $S_h$ are the mass and stiffness matrices, respectively, based on the space of continuous piecewise linear polynomials with respect to the triangulation $T_h$. Denote by the discrete total norm

$$\|||V|||_h = (2\nu\|V_\sigma\|_h^2 + 4\nu^2\|V_u\|_{1,h}^2 + \|V_p\|_h^2)^{1/2} \quad \text{with} \ V := (V_\sigma, V_u, V_p)^T.$$
4.1. Analysis of PCGM, multigrid method and a new method

The variational formulation associated with the minimization problem (3.9) can be written as: find \((\sigma^h, u^h, p^h) \in \mathcal{V}^h\) such that

\[
A_h(\sigma^h, u^h, p^h; \tau, v, q) = F_h(\tau, v, q) \quad \text{for} \quad (\tau, v, q) \in \mathcal{V}^h, \tag{4.2}
\]

where

\[
A_h(\sigma^h, u^h, p^h; \tau, v, q) = (T_h(\sqrt{2\nu} \nabla \cdot \sigma^h - \nabla p^h), \sqrt{2\nu} \nabla \cdot \tau - \nabla q) + 2\nu(\sigma^h - \sqrt{2\nu} \varepsilon(u^h), \tau - \sqrt{2\nu} \varepsilon(v)) + 4\nu^2(\nabla \cdot u^h, \nabla \cdot v) \tag{4.3}
\]

and

\[
F_h(\tau, v, q) = -(T_h f_2, \sqrt{2\nu} \nabla \cdot \tau - \nabla q) + 2\nu(f_1, \tau - \sqrt{2\nu} \varepsilon(v)) + 4\nu^2(f_3, \nabla \cdot v). \tag{4.4}
\]

Let \(\{\phi_i\}\) be the nodal basis for the space \(\mathcal{V}^h = \Phi^h \times \mathcal{U}^h \times \mathcal{P}^h\). Then the problem (4.2) can be written as the matrix problem

\[
A_h U = F, \tag{4.5}
\]

where \(A_h\) and \(F\) are given by

\[
A_h(i, j) = A_h(\phi_i, \phi_j) \quad \text{and} \quad F(i) = F_h(\phi_i).
\]

Note that the matrix \(A_h\) itself is never assembled because the operator \(T_h\) appeared in \(A_h\). Now let us set up a preconditioner for \(A_h\). Using (3.10) and (3.13) in Lemma 2 and the fact that the mass matrix is spectrally equivalent to \(h^2 I\), we can easily derive the block-diagonal matrix

\[
\text{diag}[2\nu h^2 I, 4\nu^2 S_h, h^2 I] \tag{4.6}
\]

which is spectrally equivalent to \(A_h\), where \(S_h\) is the stiffness matrix associated with (3.5). We use the inverse of the above block matrix (4.6) as a preconditioner for \(A_h\), in which the inverse of \(S_h\) is not required explicitly in the computation for the action of the preconditioner. The discrete solution operator \(T_h\) defined in (3.7) will be used instead of computing \(S_h^{-1}\). Denote by \(P_h\) a preconditioner for \(A_h\). Then, we have the following equivalent preconditioned linear algebraic system to (4.5):

\[
P_h A_h U = P_h F.
\]

We first study the performances of the preconditioner \(P_h\). From now on, we fixed \(\alpha = 0.15\). To show the effect of the preconditioner, we report the number of iterations and the condition numbers of the preconditioned linear system along with the parameter \(\beta\) in Table 1. Let \(R_m\) be the \(m\)th residual. Then

\[
(P_h R_m, R_m) \leq \varepsilon \tag{4.7}
\]

can be used to stop the iteration (see [11]). In the preconditioning conjugate gradient method, \((P_h R_m, R_m)\) is computed as part of the iteration, so the error estimator is free of cost. To reveal the
real error reduction rate and condition number of the preconditioned system, we choose $\varepsilon = 10^{-7}$. The condition number of $P_h A_h$ can be estimated by

$$\kappa(P_h A_h) \leq \left(\frac{1 + \theta}{1 - \theta}\right)^2, \quad \text{with } \theta = \left(\frac{\varepsilon}{4}\right)^{1/(2m)},$$

where $m$ is the iteration number (see [11]).

In Table 1, we present the iteration numbers and condition numbers of $P_h A_h$ for several choices of $\beta$ with the initial guess one vector. As expected, the condition numbers are uniformly bounded with respect to the grid interval $h$ and depend on $\beta$. For the case of $\beta < 0.10$ or $\beta > 2.00$, we do not report the condition numbers because we cannot reduce the relative error to $\varepsilon = 10^{-7}$ within 400 iterates, i.e., we fail to get the iteration number $m$ less than 400 to hold (4.7). From this point of view, the parameter $\beta = 0.50$ or 0.75 is a good choice for PCGM.

Now, we consider the trivial solution to (2.3), i.e., $F_1 = F_2 = F_3 = 0$. Starting with the initial guess $U^0$ of one vector, we get the $k$th iterate $U^k$ by PCGM. In fact, the $k$th iterate $U^k$ itself is the $k$th error for the trivial problem. To show the performance of PCGM, we report the convergence factors $\rho_k^\varepsilon := |||U^{k+1}|||_h/|||U^k|||_h$ defined as ratios of successive discrete total norms of errors (see Table 2).

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>1.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = \frac{1}{8}$</td>
<td>33 57.5</td>
<td>32 54.1</td>
<td>42 92.8</td>
<td>56 164.4</td>
<td>71 263.9</td>
</tr>
<tr>
<td>$h = \frac{1}{16}$</td>
<td>41 88.4</td>
<td>33 57.5</td>
<td>43 97.2</td>
<td>53 147.3</td>
<td>65 221.3</td>
</tr>
<tr>
<td>$h = \frac{1}{32}$</td>
<td>51 136.5</td>
<td>34 61.0</td>
<td>41 88.4</td>
<td>47 116.0</td>
<td>57 170.3</td>
</tr>
<tr>
<td>$h = \frac{1}{64}$</td>
<td>54 152.9</td>
<td>35 64.6</td>
<td>39 80.1</td>
<td>43 97.2</td>
<td>49 126.0</td>
</tr>
<tr>
<td>$h = \frac{1}{128}$</td>
<td>54 152.9</td>
<td>34 61.0</td>
<td>37 72.2</td>
<td>40 84.2</td>
<td>46 111.2</td>
</tr>
<tr>
<td>$h = \frac{1}{256}$</td>
<td>54 152.9</td>
<td>34 61.0</td>
<td>36 68.3</td>
<td>39 80.1</td>
<td>44 101.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>5.00</th>
<th>10.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = \frac{1}{8}$</td>
<td>0.9727</td>
<td>0.9643</td>
<td>0.9603</td>
<td>0.9617</td>
<td>0.9799</td>
<td>0.9854</td>
</tr>
<tr>
<td>$h = \frac{1}{16}$</td>
<td>0.9821</td>
<td>0.9691</td>
<td>0.9632</td>
<td>0.9637</td>
<td>0.9825</td>
<td>0.9896</td>
</tr>
<tr>
<td>$h = \frac{1}{32}$</td>
<td>0.9861</td>
<td>0.9718</td>
<td>0.9630</td>
<td>0.9648</td>
<td>0.9836</td>
<td>0.9894</td>
</tr>
<tr>
<td>$h = \frac{1}{64}$</td>
<td>0.9877</td>
<td>0.9728</td>
<td>0.9617</td>
<td>0.9639</td>
<td>0.9836</td>
<td>0.9890</td>
</tr>
<tr>
<td>$h = \frac{1}{128}$</td>
<td>0.9887</td>
<td>0.9731</td>
<td>0.9596</td>
<td>0.9617</td>
<td>0.9831</td>
<td>0.9881</td>
</tr>
<tr>
<td>$h = \frac{1}{256}$</td>
<td>0.9893</td>
<td>0.9732</td>
<td>0.9574</td>
<td>0.9581</td>
<td>0.9818</td>
<td>0.9868</td>
</tr>
</tbody>
</table>
Table 2 reports the convergence factors by PCGM measured after 50-iterations for several choices of $\beta$. This table shows that the convergence factors are independent of the grid interval $h$ but shows that $\beta = 0.75$ seems to be the best one among various choice of $\beta$. Now, we will report the convergence factors for the first some iterations and long term iterations. With fixed grid interval at $h = \frac{1}{64}$, we report the first five convergence factors and asymptotic convergence factors after 100 iterations in Table 3. This table may show that the convergence factors are worse according to iterations. The use of $\beta = 0.75$ is fairly graceful and the convergence factors of $\beta = 0.75$ is evidently better than those of $\beta = 0.25$ and 0.50. But, the larger $\beta$ is, the better the first a few convergence factors are. This means that the choice of larger $\beta$ can help the discretization error to reduce very fast during the first iterations.

In order to reflect this property, we consider a perturbed algorithm as follows. Let PCGM($\beta$, $U^0$, $k$) be the $k$th iterate by preconditioning conjugate gradient algorithm to solve the preconditioned linear system depending on $\beta$ with an initial guess $U^0$:

$$PCGM(\beta, U^0, k) \text{ to solve } P_{h,\beta}A_{h,\beta}U = P_{h,\beta}F_\beta.$$  

Since $\beta$ is a parameter used in computing $T_h = \alpha h^2 I + \beta B_h$, we do not need more cost to compute $P_{h,\beta}$, $A_{h,\beta}$ and $F_\beta$. Now we suggest a new algorithm, we call it NPCGM($\beta_1$, $\beta_2$, $k_1$, $k_2$):

**Algorithm** NPCGM($\beta_1$, $\beta_2$, $k_1$, $k_2$).

1. Set an initial guess $U^0$.
2. Find $V^1 = PCGMG(\beta_1, U^0, k_1)$.
3. Find $V^2 = PCGM(\beta_2, V^1, k_2)$.
4. Set $U^{k_1+k_2} = V^2$.

<table>
<thead>
<tr>
<th>Itr. \ $\beta$</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>5.00</th>
<th>10.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000</td>
<td>0.8778</td>
<td>0.7104</td>
<td>0.6299</td>
<td>0.5873</td>
<td>0.5041</td>
<td>0.4990</td>
</tr>
<tr>
<td>2,000</td>
<td>0.8370</td>
<td>0.8365</td>
<td>0.8246</td>
<td>0.8005</td>
<td>0.5559</td>
<td>0.5027</td>
</tr>
<tr>
<td>3,000</td>
<td>0.9460</td>
<td>0.9092</td>
<td>0.8920</td>
<td>0.8792</td>
<td>0.7695</td>
<td>0.7304</td>
</tr>
<tr>
<td>4,000</td>
<td>0.9159</td>
<td>0.9183</td>
<td>0.9043</td>
<td>0.8862</td>
<td>0.7284</td>
<td>0.6634</td>
</tr>
<tr>
<td>5,000</td>
<td>0.9574</td>
<td>0.9326</td>
<td>0.9218</td>
<td>0.9143</td>
<td>0.8338</td>
<td>0.8004</td>
</tr>
<tr>
<td>10,000</td>
<td>0.9755</td>
<td>0.9569</td>
<td>0.9385</td>
<td>0.9253</td>
<td>0.8684</td>
<td>0.8600</td>
</tr>
<tr>
<td>18,000</td>
<td>0.9821</td>
<td>0.9671</td>
<td>0.9480</td>
<td>0.9366</td>
<td>0.9560</td>
<td>0.9585</td>
</tr>
<tr>
<td>26,000</td>
<td>0.9850</td>
<td>0.9698</td>
<td>0.9531</td>
<td>0.9449</td>
<td>0.9770</td>
<td>0.9804</td>
</tr>
<tr>
<td>34,000</td>
<td>0.9864</td>
<td>0.9712</td>
<td>0.9565</td>
<td>0.9525</td>
<td>0.9809</td>
<td>0.9863</td>
</tr>
<tr>
<td>42,000</td>
<td>0.9872</td>
<td>0.9721</td>
<td>0.9592</td>
<td>0.9589</td>
<td>0.9825</td>
<td>0.9882</td>
</tr>
<tr>
<td>50,000</td>
<td>0.9877</td>
<td>0.9728</td>
<td>0.9617</td>
<td>0.9639</td>
<td>0.9836</td>
<td>0.9890</td>
</tr>
<tr>
<td>100,000</td>
<td>0.9895</td>
<td>0.9741</td>
<td>0.9694</td>
<td>0.9705</td>
<td>0.9866</td>
<td>0.9914</td>
</tr>
</tbody>
</table>
Let us again consider the problem (2.3) having the solution (4.1). To show the performance by NPCGM(β₁, β₂, k₁, k₂), we report the total iteration numbers \( k = k₁ + k₂ \) to reduce the relative error to \( ε = 10^{-7} \), i.e., satisfying (4.7), with the initial guess one in Table 4. The used grid interval is \( h = \frac{1}{64} \) and we fixed \( β₂ = 0.75 \). Table 4 shows that the use of \( β₁ = 5.0 \) or 10.0 in NPCGM(β₁, β₂, k₁, k₂) is more efficient than the use of \( β₁ = 0.75 \), i.e., the direct PCGM with \( β = 0.75 \) (see Table 1).

Finally, we discuss some numerical experiments implemented by multigrid V(1, 1)-cycle method, MGV(1, 1), to solve (4.5). The relaxation scheme we use is based on the preconditioned Richardson iteration:

\[
U^{k+1} = U^k + P_h(F - A_h U^k).
\]

First, we consider the best choice of \( α \) and \( β \) in \( T_h - α h^2 I + β B_h \) in MGV(1, 1). Tables 5 and 6 report the convergence factors by MGV(1, 1) after 20-cycles including first five cycles for a fixed \( α = 0.10 \) and \( β = 0.50 \), respectively. With a fixed \( α = 0.10 \), Table 5 shows that the use of \( β = 0.60 \) does not reduce the error after 4-cycles and the difference of the first factors between the use of \( β = 0.50 \) and 0.60 is insignificant.

Table 6 also shows that the use of \( α > 0.10 \) for a fixed \( β = 0.50 \) does not give a stable error reduction. In this point of view, the choice of \( α = 0.10 \) and \( β = 0.50 \) seems to be the best choice among the reported results. Finally, we report the iteration numbers by MGV(1, 1) to reduce the relative error to \( ε = 10^{-7} \),
i.e., satisfying (4.7), with the initial guess one in Table 7. In order to get the same relative error reduction, 32 MGV\((1,1)\)-iterations in Table 7 or 20 NPCGM-iterations in Table 4 are needed. Moreover, the cost of of MGV\((1,1)\) per one iteration is more expensive than that of NPCGM. Hence, it is clear that NPCGM is a more efficient solution method than MGV\((1,1)\) at least in our least-squares method.

Table 6
Convergence factors by MGV\((1,1)\) when \(h = \frac{1}{32}\), \(\beta = 0.5\)

<table>
<thead>
<tr>
<th>Itr. (\alpha)</th>
<th>0.05</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5810</td>
<td>0.5769</td>
<td>0.5748</td>
<td>0.5741</td>
</tr>
<tr>
<td>2</td>
<td>0.8479</td>
<td>0.8452</td>
<td>0.8409</td>
<td>0.8404</td>
</tr>
<tr>
<td>3</td>
<td>0.8773</td>
<td>0.8721</td>
<td>0.8676</td>
<td>0.9850</td>
</tr>
<tr>
<td>4</td>
<td>0.8862</td>
<td>0.8809</td>
<td>0.8764</td>
<td>2.7108</td>
</tr>
<tr>
<td>5</td>
<td>0.8904</td>
<td>0.8850</td>
<td>0.8815</td>
<td>5.3214</td>
</tr>
<tr>
<td>8</td>
<td>0.8947</td>
<td>0.8891</td>
<td>1.0438</td>
<td>5.7914</td>
</tr>
<tr>
<td>12</td>
<td>0.8960</td>
<td>0.8903</td>
<td>2.1968</td>
<td>5.8598</td>
</tr>
<tr>
<td>16</td>
<td>0.8966</td>
<td>0.8909</td>
<td>2.2928</td>
<td>5.8739</td>
</tr>
<tr>
<td>20</td>
<td>0.8970</td>
<td>0.8913</td>
<td>2.3421</td>
<td>5.8782</td>
</tr>
</tbody>
</table>

Table 7
Iteration numbers by MGV\((1,1)\) when \(\alpha = 0.10\)

<table>
<thead>
<tr>
<th>(h) (\beta)</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\frac{1}{16})</td>
<td>56</td>
<td>40</td>
<td>31</td>
</tr>
<tr>
<td>(\frac{1}{32})</td>
<td>58</td>
<td>41</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 8
Discrete \(L^2\) norm errors

<table>
<thead>
<tr>
<th>(h) (\alpha)</th>
<th>(e^h_{c_1})</th>
<th>(e^h_{c_2})</th>
<th>(e^h_{c_3})</th>
<th>(e^h_{u_1})</th>
<th>(e^h_{u_2})</th>
<th>(e^h_{p})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\frac{1}{8})</td>
<td>6.44e−1</td>
<td>5.69e−1</td>
<td>6.33e−1</td>
<td>1.43e−1</td>
<td>1.36e−1</td>
<td>2.29e−0</td>
</tr>
<tr>
<td>(\frac{1}{8})</td>
<td>2.85e−1</td>
<td>1.93e−1</td>
<td>3.00e−1</td>
<td>5.78e−2</td>
<td>5.68e−2</td>
<td>9.22e−1</td>
</tr>
<tr>
<td>(\frac{1}{16})</td>
<td>1.06e−1</td>
<td>6.40e−2</td>
<td>1.08e−1</td>
<td>1.91e−2</td>
<td>1.90e−2</td>
<td>3.29e−1</td>
</tr>
<tr>
<td>(\frac{1}{32})</td>
<td>3.33e−2</td>
<td>2.06e−2</td>
<td>3.36e−2</td>
<td>5.45e−3</td>
<td>5.44e−3</td>
<td>1.03e−1</td>
</tr>
<tr>
<td>(\frac{1}{64})</td>
<td>9.87e−3</td>
<td>6.43e−3</td>
<td>9.90e−3</td>
<td>1.43e−3</td>
<td>1.43e−3</td>
<td>3.03e−2</td>
</tr>
<tr>
<td>(\frac{1}{128})</td>
<td>2.96e−3</td>
<td>2.02e−3</td>
<td>2.96e−3</td>
<td>3.65e−4</td>
<td>3.65e−4</td>
<td>8.65e−3</td>
</tr>
<tr>
<td>(\frac{1}{256})</td>
<td>9.25e−4</td>
<td>6.54e−4</td>
<td>9.25e−4</td>
<td>9.20e−5</td>
<td>9.20e−5</td>
<td>2.50e−3</td>
</tr>
</tbody>
</table>
over components of vectors, say $E$. Discrete Tables 8 and 10 show the discretization accuracy of the finite element approximation with respect to both errors in the discrete $L^2$-norm and $H^1$-norm. Their convergence rates are reported in Tables 9 and 11, respectively.

4.2. Discretization errors

To measure the discretization error, we also consider the generalized Stokes equations (2.3) having the solution (4.1). Let $U$ be the coefficient vector of the interpolant of exact solution $(\sigma_1, \sigma_2, \sigma_3, u_1, u_2, p)$ over $V^h$. $U^h$ its approximate solution vector and let $E^h = U - U^h$. The error vector $E^h$ consists of six components of vectors, say $E^h = (e^h_{\sigma_1}, e^h_{\sigma_2}, e^h_{\sigma_3}, e^h_{u_1}, e^h_{u_2}, e^h_p)^T$. The rates of convergence for discretization errors in the discrete $L^2$- and $H^1$-norm are measured by

$$
\log_2 \frac{\|e^h_t\|_h}{\|e^{h/2}_t\|_{h/2}} \quad \text{and} \quad \log_2 \frac{\|e^h_{1,t}\|_{1,h}}{\|e^{h/2}_{1,t}\|_{1,h/2}},
$$

respectively, where $e^h_t$ denotes the $t$-components of $E^h$, i.e., $t = \sigma_1, \sigma_2, \sigma_3, u_1, u_2, p$. Numerical results in Tables 8 and 10 show the discretization accuracy of the finite element approximation with respect to both discrete $L^2$-norm and $H^1$-norm. Their convergence rates are reported in Tables 9 and 11, respectively. The theoretically predicted error bounds are only $O(h)$ in $L^2$ for the stress and pressure and $O(h)$ in $H^1$ for the velocity. The resulting convergence rates for the velocity asymptotically approach to the best approximation rates $O(h^2)$ in $L^2$ and $O(h^{\sqrt{2}+\delta})$ in $H^1$. It appears that we obtained optimal convergence

### Table 9
Discrete $L^2$ norm convergence rates

<table>
<thead>
<tr>
<th>$h$</th>
<th>$e^h_{\sigma_1}$</th>
<th>$e^h_{\sigma_2}$</th>
<th>$e^h_{\sigma_3}$</th>
<th>$e^h_{u_1}$</th>
<th>$e^h_{u_2}$</th>
<th>$e^h_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{8}$</td>
<td>1.1736</td>
<td>1.5584</td>
<td>1.0770</td>
<td>1.3091</td>
<td>1.2584</td>
<td>1.3133</td>
</tr>
<tr>
<td>$\frac{1}{16}$</td>
<td>1.4297</td>
<td>1.5941</td>
<td>1.4728</td>
<td>1.5975</td>
<td>1.5820</td>
<td>1.4835</td>
</tr>
<tr>
<td>$\frac{1}{32}$</td>
<td>1.6680</td>
<td>1.6362</td>
<td>1.6867</td>
<td>1.8082</td>
<td>1.8034</td>
<td>1.6665</td>
</tr>
<tr>
<td>$\frac{1}{64}$</td>
<td>1.7569</td>
<td>1.6802</td>
<td>1.7632</td>
<td>1.9265</td>
<td>1.9246</td>
<td>1.7764</td>
</tr>
<tr>
<td>$\frac{1}{128}$</td>
<td>1.7385</td>
<td>1.6690</td>
<td>1.7410</td>
<td>1.9751</td>
<td>1.9741</td>
<td>1.8081</td>
</tr>
<tr>
<td>$\frac{1}{256}$</td>
<td>1.6781</td>
<td>1.6284</td>
<td>1.6795</td>
<td>1.9882</td>
<td>1.9882</td>
<td>1.7898</td>
</tr>
</tbody>
</table>

### Table 10
Discrete $H^1$ norm errors

<table>
<thead>
<tr>
<th>$h$</th>
<th>$e^h_{\sigma_1}$</th>
<th>$e^h_{\sigma_2}$</th>
<th>$e^h_{\sigma_3}$</th>
<th>$e^h_{u_1}$</th>
<th>$e^h_{u_2}$</th>
<th>$e^h_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{4}$</td>
<td>3.26e+0</td>
<td>4.17e+0</td>
<td>3.95e+0</td>
<td>7.05e-1</td>
<td>7.05e-1</td>
<td>1.75e+1</td>
</tr>
<tr>
<td>$\frac{1}{8}$</td>
<td>2.38e+0</td>
<td>1.74e+0</td>
<td>2.90e+0</td>
<td>2.88e-1</td>
<td>2.96e-1</td>
<td>7.48e+0</td>
</tr>
<tr>
<td>$\frac{1}{16}$</td>
<td>1.64e+0</td>
<td>9.37e-1</td>
<td>1.76e+0</td>
<td>1.03e-1</td>
<td>1.05e-1</td>
<td>3.78e+0</td>
</tr>
<tr>
<td>$\frac{1}{32}$</td>
<td>1.12e+0</td>
<td>6.02e-1</td>
<td>1.14e+0</td>
<td>3.20e-2</td>
<td>3.22e-2</td>
<td>2.36e+0</td>
</tr>
<tr>
<td>$\frac{1}{64}$</td>
<td>7.88e-1</td>
<td>4.04e-1</td>
<td>7.92e-1</td>
<td>9.16e-3</td>
<td>9.18e-3</td>
<td>1.61e+0</td>
</tr>
<tr>
<td>$\frac{1}{128}$</td>
<td>5.57e-1</td>
<td>2.78e-1</td>
<td>5.57e-1</td>
<td>2.61e-3</td>
<td>2.61e-3</td>
<td>1.12e+0</td>
</tr>
<tr>
<td>$\frac{1}{256}$</td>
<td>3.93e-1</td>
<td>1.94e-1</td>
<td>3.94e-1</td>
<td>7.74e-4</td>
<td>7.74e-4</td>
<td>7.92e-1</td>
</tr>
</tbody>
</table>
Table 11
Discrete $H^1$-norm convergence rates

<table>
<thead>
<tr>
<th>$h$</th>
<th>$e_{u_1}^h$</th>
<th>$e_{u_2}^h$</th>
<th>$e_{u_3}^h$</th>
<th>$e_{u_1}^h$</th>
<th>$e_{u_2}^h$</th>
<th>$e_p^h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = \frac{1}{8}$</td>
<td>0.4536</td>
<td>1.2609</td>
<td>0.4450</td>
<td>1.2898</td>
<td>1.2514</td>
<td>1.2282</td>
</tr>
<tr>
<td>$h = \frac{1}{16}$</td>
<td>0.5312</td>
<td>0.8936</td>
<td>0.7211</td>
<td>1.4766</td>
<td>1.4954</td>
<td>0.9817</td>
</tr>
<tr>
<td>$h = \frac{1}{32}$</td>
<td>0.5496</td>
<td>0.6369</td>
<td>0.6198</td>
<td>1.6921</td>
<td>1.7034</td>
<td>0.6789</td>
</tr>
<tr>
<td>$h = \frac{1}{64}$</td>
<td>0.5141</td>
<td>0.5756</td>
<td>0.5329</td>
<td>1.8076</td>
<td>1.8120</td>
<td>0.5526</td>
</tr>
<tr>
<td>$h = \frac{1}{128}$</td>
<td>0.5018</td>
<td>0.5401</td>
<td>0.5066</td>
<td>1.8104</td>
<td>1.8122</td>
<td>0.5181</td>
</tr>
<tr>
<td>$h = \frac{1}{256}$</td>
<td>0.4998</td>
<td>0.5196</td>
<td>0.5010</td>
<td>1.7553</td>
<td>1.7570</td>
<td>0.5078</td>
</tr>
</tbody>
</table>

in the $L^2$-norm and super-convergence in the $H^1$-norm for the velocity component. On the other hand, the convergence rates for the stress and pressure are like $O(h^{3/2+\delta})$ in $L^2$ and $O(h^{1/2})$ in $H^1$. The computed error bounds for the stress and pressure are better than those of the theoretically predicted, but they are suboptimal in comparison with the best approximation $O(h^2)$ in $L^2$ and $O(h)$ in $H^1$. In order to obtain the optimal convergence rates for all unknowns, one may choose the polynomials of one degree lower for the stress and pressure than that for the velocity. However, the use of a single approximating space for all variables simplifies programming of least-squares finite element methods.

Acknowledgements

The authors would like to thank the anonymous referees for their valuable suggestions and comments.

References