SIAM J. NUMER. ANAL.
Vol. 47, No. 6, pp. 4098–4111

OPTIMAL ERROR ESTIMATE FOR THE DIV LEAST-SQUARES METHOD WITH DATA $f \in L^2$ AND APPLICATION TO NONLINEAR PROBLEMS*

ZHIQIANG CAI† AND JAEUN KU‡

Abstract. The div least-squares methods have been studied by many researchers for the second-order elliptic equations, elasticity, and the Stokes equations, and optimal error estimates have been obtained in the $H(\text{div}) \times H^1$ norm. However, there is no known convergence rate when the given data f belongs only to L^2 space. In this paper, we will establish an optimal error estimate in the $L^2 \times H^1$ norm with the given data $f \in L^2$ and, hence, fill a theoretical gap of least-squares methods. As a consequence of this estimate, we will provide a convergence analysis for the linearization process on solving Navier–Stokes equations, which uses the div least-squares method for solving the corresponding Stokes equations.

Key words. least-squares method, error estimate, elliptic equations, elasticity, Stokes, incompressible Newtonian flow

AMS subject classifications. 65M60, 65M15

DOI. 10.1137/080738350

1. Introduction. There is substantial interest in the use of least-squares principles for the approximate solution of partial differential equations with applications in both solid and fluid mechanics. Loosely speaking, there are three types of least-squares methods: the inverse approach, the div approach, and the div-curl approach. For the least-squares methods such as the inverse approach and div-curl approach, we refer to [1, 4, 7, 11, 24, 27] and references therein.

For the scalar elliptic equations, the div approach based on the flux-pressure (the respective dual and primal variables) formulation has been studied by many researchers (see, e.g., [5, 10, 14, 15, 17, 18, 20, 21, 22, 23, 25, 26]). For elasticity and the Stokes equations, the div least-squares approach based on the stress-displacement and the stress-velocity (the respective dual and primal variables) formulations, respectively, was introduced and analyzed in [12, 13]. An optimal error estimate is obtained in $H(\text{div}) \times H^1$ norm. This estimate also yields an optimal L^2 norm error estimate for the dual variable. Recently, we in [9] established the optimal L^2 estimate of the primal variable. But, all these estimates mentioned above are under the assumption that the right-hand side function f is as smooth as the divergence of the dual variable, which is slightly more than those of the Galerkin finite element method. In particular, when the given date f belongs only to L^2 space, there is no known convergence rate. This, in turn, causes some difficulties for a convergence analysis for the linearization process for solving nonlinear problems such as Navier–Stokes equations by using the abstract theory of Brezzi, Rappaz, and Raviart (see, e.g., [19]).

One of the purposes of this paper is to study error estimates on the div least-squares approach when f is only in L^2 and, hence, to fill a theoretical gap of the

*Received by the editors October 17, 2008; accepted for publication (in revised form) September 22, 2009; published electronically January 15, 2010.
http://www.siam.org/journals/sinum/47-6/73835.html

†Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067 (zcai@math.purdue.edu). This author's research was supported in part by the National Science Foundation under grant DMS-0511430.

‡Department of Mathematics, Oklahoma State University, Mathematical Sciences 401, Stillwater, OK 74078 (jku@math.okstate.edu).

div least-squares method. The other is to analyze the div least-squares finite element method for the stationary Navier–Stokes equations. When $f \in L^2$, it is obvious that the divergence of the dual variable belongs only to the L^2 space. Hence, it is impossible to obtain a convergence rate in the $H(\text{div}) \times H^1$ norm. Instead, in this paper, we will establish linear rate convergence in a weaker norm, the $L^2 \times H^1$ norm for the respective dual and primal variables. This will be done by modifying the argument of Yang in [28]. Note that this is not a trivial task simply because the corresponding "energy" norm of the div least-squares functional differs from the $L^2 \times H^1$ norm. This new estimate plays a key role in convergence analysis for the least-squares finite element method for the stationary Navier–Stokes equations, which will be done by verifying various assumptions of the abstract theory by Brezzi, Rappaz, and Raviart.

Previous least-squares approaches to solve nonlinear problems such as Navier–Stokes equations are based mostly on the div-curl approach [2, 3]. This type of approach requires the space for dual variable in H^1 space and its approximate space to be C^0 elements, and thus prevent using classical mixed elements such as Raviart–Thomas elements. However, it is reported [6] that RT elements perform favorably over the standard C^0 element when locally conservative approximation is essential.

The paper is organized as follows. Section 2 introduces mathematical equations for the second-order scalar elliptic partial differential equations, elasticity, and the Stokes equations. The div least-squares formulation and its finite element approximation for those equations are described in section 3. We want to note that sections 2 and 3 are a brief summary, and we refer the reader to [9] for a detailed presentation. In section 4, we establish the optimal L^2 error estimates for the flux and optimal H^1 error estimates for the primary variable with minimum regularity assumption. The optimal H^1 error estimates play the key role in a convergence analysis for the linearization process to solve Navier–Stokes equations in section 5.

2. Mathematical equations.

Let Ω be a bounded, open, connected subset of $\Re^d$ ($d = 2$ or 3) with a Lipschitz continuous boundary $\partial\Omega$, where $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. For simplicity, we assume that Γ_D is not empty (i.e., $\text{mes}\,(\Gamma_D) \neq 0$).

We use the standard notations and definitions for the Sobolev spaces $H^s(\Omega)^d$ for $s \geq 0$. Set

$$H^1_D(\Omega) := \{q \in H^1(\Omega) \, : \, q = 0 \text{ on } \Gamma_D\}.$$

When $\Gamma = \partial\Omega$, denote $H^1_D(\Omega)$ by $H^1_0(\Omega)$. Let

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

which is a Hilbert space under the norm $\|\mathbf{v}\|_{H(\text{div};\,\Omega)} = (\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2)^{\frac{1}{2}}$, and define the subspace

$$H_N(\text{div};\Omega) = \{\mathbf{v} \in H(\text{div};\Omega) \, : \, \mathbf{n} \cdot \mathbf{v} = 0 \text{ on } \Gamma_N\},$$

where $\mathbf{n}$ is the outward unit vector normal to the boundary.

2.1. Second-order elliptic problems.

Consider the following second-order elliptic boundary value problem:

$$-\nabla \cdot (A\nabla u) + Xu = f \quad \text{in } \Omega \tag{2.1}$$

with boundary conditions

$$u = 0 \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n} \cdot A\nabla u = 0 \quad \text{on } \Gamma_N, \tag{2.2}$$

where A is a given $d \times d$ tensor function, uniformly symmetric positive definite; $X\,u = \mathbf{b} \cdot \nabla u + c\,u$ with $\mathbf{b}, c$ smooth functions; and f is given scalar function. For simplicity of the presentation, assume that (2.1) satisfies the full H^2 regularity estimates:

$$\|u\|_2 \le C\,\|f\|_0. \tag{2.3}$$

2.2. Elasticity and Stokes equations. Let

$$\boldsymbol{\epsilon}(\mathbf{v}) = \big(\epsilon_{ij}(\mathbf{v})\big)_{d\times d} \quad \text{with} \quad \epsilon_{ij}(\mathbf{v}) = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right).$$

For a tensor function $\boldsymbol{\tau} = (\tau_{ij})_{d\times d}$, let $\boldsymbol{\tau}_i = (\tau_{i1}, \ldots, \tau_{id})$ denote its i^{th}-row for $i = 1, \ldots, d$ and define its divergence, normal, and trace by

$$\nabla \cdot \boldsymbol{\tau} = (\nabla \cdot \boldsymbol{\tau}_1, \ldots, \nabla \cdot \boldsymbol{\tau}_d), \quad \mathbf{n} \cdot \boldsymbol{\tau} = (\mathbf{n} \cdot \boldsymbol{\tau}_1, \ldots, \mathbf{n} \cdot \boldsymbol{\tau}_d), \quad \text{and} \quad \operatorname{tr} \boldsymbol{\tau} = \sum_{i=1}^{d} \tau_{ii},$$

respectively.

Elasticity and Stokes equations may be casted as the following pressure-perturbed form of the generalized Stokes equations

$$\begin{cases} -\nabla \cdot (2\mu\,\boldsymbol{\epsilon}(\mathbf{u}) - p\,\boldsymbol{\delta}) = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} + \dfrac{1}{\lambda}\,p = g & \text{in } \Omega \end{cases} \tag{2.4}$$

with boundary conditions

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n} \cdot (2\mu\,\boldsymbol{\epsilon}(\mathbf{u}) - p\,\boldsymbol{\delta}) = \mathbf{0} \quad \text{on } \Gamma_N. \tag{2.5}$$

When $\Gamma_N = \emptyset$, p is unique up to an additive constant provided that the compatibility condition $\int_\Omega g\,dx = 0$ holds. Again, we assume that (2.4) satisfies the full H^2 regularity estimates:

$$\begin{cases} \|\mathbf{u}\|_2 + \|p\|_1 \le C\,(\|\mathbf{f}\|_0 + \|g\|_1) & \text{for } \lambda = \infty, \\ \|\mathbf{u}\|_2 + \lambda\,\|\nabla \cdot \mathbf{u}\|_1 \le C\,\|\mathbf{f}\|_0 & \text{for } \lambda < \infty, \end{cases} \tag{2.6}$$

where C is a positive constant independent of $\mathbf{u}$, p, and λ.

3. Least-squares formulations.

3.1. First-order systems. Introducing the flux (vector) variable $\sigma = -A\,\nabla u$, the scalar elliptic problem in (2.1) may be rewritten as the following first-order partial differential system

$$\begin{cases} A^{-1}\sigma + \nabla u = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot \sigma + X u = f & \text{in } \Omega \end{cases} \tag{3.1}$$

with boundary conditions

$$u = 0 \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n} \cdot \sigma = 0 \quad \text{on } \Gamma_N. \tag{3.2}$$

For the generalized Stokes equations, define $\mathcal{A}_\lambda$ by

$$\mathcal{A}_\lambda\,\boldsymbol{\tau} = \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{\lambda}{d\lambda + 2\mu}(\operatorname{tr}\boldsymbol{\tau})\,\boldsymbol{\delta}\right) \quad \text{for } \lambda \in (0, \infty]. \tag{3.3}$$

Without loss of generality, we take $\mu = 1/2$. Note that $\frac{\lambda}{d\lambda+2\mu} = \frac{1}{d}$ when $\lambda = \infty$. Consider the following first-order system of partial differential equations corresponding to (2.4) studied in [13, 12]:

$$\begin{cases} \mathcal{A}_\lambda \boldsymbol{\sigma} - \boldsymbol{\epsilon}(\mathbf{u}) = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{\sigma} = -\mathbf{f} & \text{in } \Omega \end{cases} \tag{3.4}$$

with boundary conditions

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n} \cdot \boldsymbol{\sigma} = \mathbf{0} \quad \text{on } \Gamma_N. \tag{3.5}$$

3.2. Least-squares variational problems. For the solution spaces, define

$$\mathbf{X}_\lambda \equiv \begin{cases} H_N(\text{div};\,\Omega)^d & \text{for } \lambda \in (0,\,\infty), \\ \{\boldsymbol{\tau} \in H_N(\text{div};\,\Omega)^d \mid \int_\Omega \operatorname{tr} \boldsymbol{\tau}\, dx = 0\} & \text{for } \lambda = \infty, \end{cases}$$

and let

$$\Sigma = \begin{cases} H_N(\text{div};\Omega) & \text{for (3.1)}, \\ \mathbf{X}_\lambda & \text{for (3.4)} \end{cases} \quad \text{and} \quad U = \begin{cases} H^1_D(\Omega) & \text{for (3.1)}, \\ H^1_D(\Omega)^d & \text{for (3.4)}. \end{cases}$$

Define a bilinear form $b(\cdot;\,\cdot)$ on $(\Sigma \times U) \times (\Sigma \times U)$ by

$$b(\boldsymbol{\sigma}, \mathbf{u};\, \boldsymbol{\tau}, \mathbf{v}) = \begin{cases} (\sigma + A\nabla u,\, A^{-1}\tau + \nabla v) + (\nabla \cdot \sigma + Xu,\, \nabla \cdot \tau + Xv) & \text{for (3.1)}, \\ (\mathcal{A}_\lambda \boldsymbol{\sigma} - \boldsymbol{\epsilon}(\mathbf{u}),\, \mathcal{A}_\lambda \boldsymbol{\tau} - \boldsymbol{\epsilon}(\mathbf{v})) + (\nabla \cdot \boldsymbol{\sigma},\, \nabla \cdot \boldsymbol{\tau}) & \text{for (3.4)} \end{cases}$$

for all $(\boldsymbol{\sigma},\, \mathbf{u};\, \boldsymbol{\tau},\, \mathbf{v}) \in (\Sigma \times U) \times (\Sigma \times U)$ and a linear form $F(\cdot)$ on $\Sigma \times U$ by

$$F(\boldsymbol{\tau}, \mathbf{v}) = \begin{cases} (f,\, \nabla \cdot \tau + Xv) & \text{for (3.1)}, \\ (-\mathbf{f},\, \nabla \cdot \boldsymbol{\tau}) & \text{for (3.4)} \end{cases}$$

for all $(\boldsymbol{\tau},\, \mathbf{v}) \in \Sigma \times U$. Then the true solution $(\boldsymbol{\sigma},\, \mathbf{u}) \in \Sigma \times U$ satisfies

$$b(\boldsymbol{\sigma},\, \mathbf{u};\, \boldsymbol{\tau},\, \mathbf{v}) = F(\boldsymbol{\tau},\, \mathbf{v}) \quad \forall\, (\boldsymbol{\tau},\, \mathbf{v}) \in \Sigma \times U. \tag{3.6}$$

3.3. Least-squares finite element approximations. For simplicity of presentation, we consider only triangular and tetrahedra elements for the respective two and three dimensions. Assuming that the domain Ω is polygonal, let $\mathcal{T}_h$ be a regular triangulation of Ω (see [16]) with triangular/tetrahedra elements. Let $P_k(K)$ be the space of polynomials of degree k on triangle K and $RT_k(K)$ denote the local Raviart–Thomas space of order k on K (see [19]). Then the standard $H(\text{div};\,\Omega)$ conforming Raviart–Thomas space of index k [8] and the standard (conforming) continuous piecewise polynomials of degree $k+1$ are defined, respectively, by

$$\Sigma_h^k = \{\boldsymbol{\tau} \in \Sigma \,:\, \boldsymbol{\tau}|_K \in RT_k(K)^m \quad \forall\, K \in \mathcal{T}_h\}, \tag{3.7}$$

$$V_h^{k+1} = \{\mathbf{v} \in U \quad \mathbf{v}|_K \in P_{k+1}(K)^m \quad \forall\, K \in \mathcal{T}_h\},$$

where $m = 1$ or $m = d$. It is well-known (see [16]) that V_h^{k+1} has the following approximation property: let $k \geq 0$ be an integer and let $l \in [0,\, k+1]$

$$\inf_{\mathbf{u} \in V_h^{k+1}} \|\mathbf{u} - \mathbf{v}\|_1 \leq Ch^l\, \|\mathbf{u}\|_{l+1}, \tag{3.8}$$

for $\mathbf{u} \in H^{l+1}(\Omega)^m \cap U$. It is also well-known (see [8]) that Σ_h^k has the following approximation property: let $k \geq 0$ be an integer and let $l \in [1, k+1]$

$$\inf_{\boldsymbol{\tau} \in \Sigma_h^k} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{H(\mathrm{div};\,\Omega)} \leq Ch^l \left(\|\boldsymbol{\sigma}\|_l + \|\nabla \cdot \boldsymbol{\sigma}\|_l\right) \tag{3.9}$$

for $\boldsymbol{\sigma} \in H^l(\Omega)^{m\times m} \cap \Sigma$ with $\nabla \cdot \boldsymbol{\sigma} \in H^l(\Omega)^m$.

The finite element discretization of our least-squares variational problem is: find $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \Sigma_h^k \times V_h^{k+1}$ such that

$$b(\boldsymbol{\sigma}_h, \mathbf{u}_h; \boldsymbol{\tau}, \mathbf{v}) = F(\boldsymbol{\tau}, \mathbf{v}) \quad \forall\, (\boldsymbol{\tau}, \mathbf{v}) \in \Sigma_h^k \times V_h^{k+1}. \tag{3.10}$$

Note that we have the following orthogonality property:

$$b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h; \boldsymbol{\tau}, \mathbf{v}) = 0 \quad \forall\, (\boldsymbol{\tau}, \mathbf{v}) \in \Sigma_h^k \times V_h^{k+1}. \tag{3.11}$$

4. Error estimates with $\boldsymbol{f} \in \boldsymbol{L^2}$. In this section, we take $k = 0$ in (3.7). The extension for the case $k \geq 1$ is straightforward. We will use the operator $\pi_h : \Sigma \to \Sigma_h^1$, see [8, III and IV] for $m = 1$ (i.e., the scalar case) with the following property:

$$\|\boldsymbol{\sigma} - \pi_h \boldsymbol{\sigma}\|_0 \leq Ch\|\boldsymbol{\sigma}\|_1 \quad \forall\, \boldsymbol{\sigma} \in H^1(\Omega)^d, \tag{4.1}$$

$$(\nabla \cdot (\boldsymbol{\sigma} - \pi_h \boldsymbol{\sigma}), q) = 0 \quad \forall\, q \in Q_h, \tag{4.2}$$

where

$$Q_h = \{q \in L^2(\Omega) \,:\, q_{|\kappa} = \mathit{constant} \quad \text{for each } \kappa \in \mathcal{T}_h\}. \tag{4.3}$$

By using the above operator and modifying the argument in [28], we obtain our main results for least-squares methods with the minimum assumption, i.e., $f \in L^2$. It turns out that the proofs for the second-order scalar elliptic problems are more complicated than the ones for Stokes equations and linear elasticity problems due to the lower-order terms in (2.1). Thus, we will provide the proof for the scalar case first in detail and provide the sketch of the proof for the linear elasticity and Stokes equations.

4.1. The scalar elliptic problems.

THEOREM 4.1. *Let* $(\boldsymbol{\sigma}, u)$ *and* $(\boldsymbol{\sigma}_h, u_h)$ *be the solutions of* (3.6) *corresponding to the scalar problems* (3.1) *and* (3.10). *Then, the following a priori estimate holds:*

$$\|u - u_h\|_1 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \leq Ch(\|u\|_2 + \|\boldsymbol{\sigma}\|_1) \leq Ch\|f\|_0.$$

Proof. By the triangle inequality, we have

$$\|u - u_h\|_1^2 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0^2 \leq C(\|u - u_I\|_1^2 + \|\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma}\|_0^2 + \|u_I - u_h\|_1^2 + \|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0^2).$$

Let u_I be the interpolation of u, which satisfies the approximation property in (3.8). Using (4.1) along with u_I, we have

$$\|u - u_I\|_1^2 + \|\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma}\|_0^2 \leq Ch^2(\|u\|_2^2 + \|\boldsymbol{\sigma}\|_1^2).$$

Thus, in order to prove the theorem, it suffices to prove that

$$\|u_I - u_h\|_1^2 + \|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0^2 \leq Ch^2(\|u\|_2^2 + \|\boldsymbol{\sigma}\|_1^2). \tag{4.4}$$

By the coercivity of the least-squares functional and the orthogonality property of least-squares solution $(\boldsymbol{\sigma}_h, u_h)$, we have

$$
\begin{aligned}
C\left(\|u_I - u_h\|_1^2 + \|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}^2\right) &\le b(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_I - u_h; \pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_I - u_h) \\
&= b(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}, u_I - u; \pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_I - u_h) \qquad (4.5)\\
&= I_1 + I_2 + I_3
\end{aligned}
$$

with

$$
I_1 = (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), \nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)), \qquad I_2 = (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), X(u_I - u_h)),
$$

$$
\begin{aligned}
\text{and } I_3 &= (X(u_I - u), \nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + X(u_I - u_h)) \\
&\quad + ((\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}) + \mathcal{A}\nabla(u_I - u), \mathcal{A}^{-1}(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + \nabla(u_I - u_h)).
\end{aligned}
$$

Using (4.2) and the fact that $\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \in Q_h$ yields

$$
I_1 = 0.
$$

Let $\mathbf{b}_I$ be a piecewise constant function such that

$$
\|\mathbf{b} - \mathbf{b}_I\|_{L_\infty} \le Ch\|\mathbf{b}\|_{W^1_\infty} \le Ch. \qquad (4.6)
$$

Since $Xv = \mathbf{b}\cdot\nabla v + cv$, then

$$
\begin{aligned}
I_2 &= (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), \mathbf{b}\cdot\nabla(u_I - u_h)) + (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), c(u_I - u_h)) \\
&= (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), (\mathbf{b} - \mathbf{b}_I)\cdot\nabla(u_I - u_h)) + (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), c(u_I - u_h)).
\end{aligned}
$$

The last equality used the identity

$$
(\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), \mathbf{b}_I\cdot\nabla(u_I - u_h)) = 0,
$$

which follows from (4.2) and the fact that $\mathbf{b}_I\cdot\nabla(u_I - u_h) \in Q_h$. Thus, by the Cauchy–Schwarz inequality, (4.6), integration by parts, the arithmetic-geometric inequality, and (4.1), we have

$$
\begin{aligned}
I_2 &\le \|\nabla\cdot(\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma})\|_0\, \|(\mathbf{b} - \mathbf{b}_I)\cdot\nabla(u_I - u_h)\|_0 + (\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}), c(u_I - u_h)) \\
&\le C\|\mathbf{b} - \mathbf{b}_I\|_{L_\infty}\, \|\nabla\cdot(\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma})\|_0\, \|u_I - u_h\|_1 + (\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}, \nabla(c(u_I - u_h))) \\
&\le Ch\|\nabla\cdot(\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma})\|_0\, \|u_I - u_h\|_1 + \|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_0\, \|\nabla(c(u_I - u_h))\|_0 \\
&\le C\left(\|\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma}\|_0 + h\,\|\nabla\cdot(\boldsymbol{\sigma} - \pi_h\boldsymbol{\sigma})\|_0\right)\, \|u_I - u_h\|_1 \\
&\le Ch\|\boldsymbol{\sigma}\|_1\, \|u_I - u_h\|_1.
\end{aligned}
$$

For I_3, it follows from the Cauchy–Schwarz inequality, the triangle inequality, the approximation properties of u_I and $\pi_h\boldsymbol{\sigma}$, and the arithmetic-geometric inequality that

$$
\begin{aligned}
I_3 &\le \|u - u_I\|_1\left(\|\nabla\cdot(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|u_I - u_h\|_1\right) \\
&\quad + \left(\|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_0 + \|u - u_I\|_1\right)\left(\|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|u_I - u_h\|_1\right) \\
&\le Ch\left(\|u\|_2 + \|\boldsymbol{\sigma}\|_1\right)\left(\|u_I - u_h\|_1 + \|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}\right).
\end{aligned}
$$

Combining (4.5) with bounds for I_2 and I_3 gives

$$\|u_I - u_h\|_1^2 + \|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}^2 \le Ch\left(\|u\|_2 + \|\boldsymbol{\sigma}\|_1\right)\left(\|u_I - u_h\|_1 + \|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}\right),$$

which implies (4.4). Finally, by the a priori estimate in (2.3) and using the fact that $\|\boldsymbol{\sigma}\|_1 < \|u\|_2$, we complete the proof of the theorem. ∎

4.2. The generalized Stokes equations. We now state the corresponding result for linear elasticity and Stokes equations. Here, we use the following operator:

$$\boldsymbol{\pi}_h \boldsymbol{\sigma} = (\pi_h \boldsymbol{\sigma}_1, \dots, \pi_h \boldsymbol{\sigma}_d), \tag{4.7}$$

where π_h is operator satisfying (4.1) and (4.2). Note that $\boldsymbol{\pi}_h$ satisfies the following properties:

$$\|\boldsymbol{\sigma} - \boldsymbol{\pi}_h \boldsymbol{\sigma}\|_0 \le Ch\|\boldsymbol{\sigma}\|_1 \quad \forall\, \boldsymbol{\sigma} \in H^1(\Omega)^{d\times d}, \tag{4.8}$$

$$(\nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\pi}_h \boldsymbol{\sigma}), \mathbf{q}) = 0 \quad \forall\, \mathbf{q} \in Q_h^d. \tag{4.9}$$

THEOREM 4.2. *Let* $(\boldsymbol{\sigma}, \mathbf{u})$ *and* $(\boldsymbol{\sigma}_h, \mathbf{u}_h)$ *be the solutions of* (3.6) *corresponding to the generalized Stokes equations* (3.3) *and* (3.10). *Then, the following a priori error estimate holds:*

$$\|\mathbf{u} - \mathbf{u}_h\|_1 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \le Ch(\|\mathbf{u}\|_2 + \|\boldsymbol{\sigma}\|_1) \le Ch\|\mathbf{f}\|_0.$$

Proof. The theorem may be proved in a similar fashion as that of Theorem 4.1. For the convenience of readers, we sketch the proof here. To this end, again it suffices to show that

$$\|\mathbf{u}_I - \mathbf{u}_h\|_1^2 + \|\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0^2 \le Ch^2(\|\mathbf{u}\|_2^2 + \|\boldsymbol{\sigma}\|_1^2). \tag{4.10}$$

By the coercivity of the least-squares functional and the orthogonality property of least-squares solution $(\boldsymbol{\sigma}_h, u_h)$, we have

$$\begin{aligned} C\left(\|\mathbf{u}_I - \mathbf{u}_h\|_1^2 + \|\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}^2\right) &\le b(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h; \boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h) \\ &= b(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}, \mathbf{u}_I - \mathbf{u}; \boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h) \\ &= I_1 + I_2 \end{aligned} \tag{4.11}$$

with

$$I_1 = (\nabla\cdot(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\nabla\cdot(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h))$$

$$\text{and} \quad I_2 = (\mathcal{A}_\lambda(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}) - \boldsymbol{\epsilon}(\mathbf{u}_I - \mathbf{u}), \quad \mathcal{A}_\lambda(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \boldsymbol{\epsilon}(\mathbf{u}_I - \mathbf{u}_h)).$$

Using (4.9) and the fact that $\nabla\cdot(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \in Q_h^d$ yields

$$I_1 = 0.$$

It follows from the Cauchy–Schwarz inequality, the definition of $\mathcal{A}_\lambda$, and the approximation properties (4.8) and (3.8) that

$$\begin{aligned} I_2 &\le \left(\|\mathcal{A}_\lambda(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_0 + \|\mathbf{u}_I - \mathbf{u}\|_1\right)\left(\|\mathcal{A}_\lambda(\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|\mathbf{u}_I - \mathbf{u}_h\|_1\right) \\ &\le \left(\|\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_0 + \|\mathbf{u}_I - \mathbf{u}\|_1\right)\left(\|\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|\mathbf{u}_I - \mathbf{u}_h\|_1\right) \\ &\le Ch\left(\|\mathbf{u}\|_2 + \|\boldsymbol{\sigma}\|_1\right)\left(\|\boldsymbol{\pi}_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}^2 + \|\mathbf{u}_I - \mathbf{u}_h\|_1^2\right)^{1/2}. \end{aligned}$$

Combining the results for I_1 and I_2 and (4.11) implies the validity of (4.10). This completes the proof of the theorem. ∎

5. Application to nonlinear problems.

5.1. Problem formulation. In this section, we will use the error estimate developed earlier to provide a convergence analysis to a linearization process for solving Navier–Stokes equation of the form:

$$\begin{cases} -\nu \triangle \mathbf{u} + \sum_{j=1}^{d} u_j \partial \mathbf{u}/\partial x_j + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.1}$$

By taking $\boldsymbol{\sigma} = \nu\nabla\mathbf{u} - p\,I$ and rescaling the stress, the pressure, and the right-hand side by $\boldsymbol{\sigma}/\nu \to \boldsymbol{\sigma}$, $p/\nu \to p$ and $\mathbf{f}/\nu \to \mathbf{f}$, respectively, we obtain the following first-order system:

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma} + \dfrac{1}{\nu} \displaystyle\sum_{j=1}^{d} u_j \partial \mathbf{u}/\partial x_j = \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma} - \nabla \mathbf{u} + p\,I = 0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.2}$$

In order to solve the above nonlinear problem, we closely follow the abstract framework of [19]. Let

$$\mathbf{X} = H(div)^d \times \left(W_2^{\frac{3}{2}-\delta}\right)^d \times L_0^2 \quad \text{and} \quad \mathbf{Y} = L^2, \tag{5.3}$$

where $L_0^2 = \{q \in L^2 : \int_\Omega q = 0\}$, $\delta > 0$ is sufficiently small, and $d = 2$ or 3. We want to point out that we choose $(W_2^{\frac{3}{2}-\delta})^d$ space, not $(W_2^1)^d$ for the space of $\mathbf{u}$ in the definition of $\mathbf{X}$. This is necessary to have the nonlinear term $\sum_{j=1}^{d} u_j \partial \mathbf{u}/\partial x_j$ to be in L^2 space. Note also that $V_h^1 \subset (W_2^{\frac{3}{2}-\delta})^d$, where V_h^1 is defined in (3.7).

Consider the nonlinear problem of the form

$$F(\lambda, \mathcal{U}) \equiv \mathcal{U} + TG(\lambda, \mathcal{U}) = 0 \ \text{ for } \ \mathcal{U} = (\boldsymbol{\sigma}, \mathbf{u}, p) \in \mathbf{X} \ \text{ and } \ \lambda \in \Lambda, \tag{5.4}$$

where $\Lambda \subset (0, +\infty)$ is a compact interval. Define a linear operator T as follows: given $\mathbf{f}_* \in \mathbf{Y}$, denote by

$$(\boldsymbol{\sigma}_*, \mathbf{u}_*, p_*) = T\mathbf{f}_* \in \mathbf{X}$$

the solution of the Dirichlet problem for the Stokes equations

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma}_* = \mathbf{f}_* & \text{in } \Omega, \\ \boldsymbol{\sigma}_* - \nabla \mathbf{u}_* + p_* I = 0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{u}_* = 0 & \text{in } \Omega, \\ \mathbf{u}_* = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.5}$$

Also, with data $\mathbf{f}_* \in \mathbf{Y}$, associate $\mathcal{C}^\infty$-mapping from $\mathbb{R}_+ \times \mathbf{X}$ into $\mathbf{Y}$ defined by

$$G : (\lambda, \mathcal{V} = (\boldsymbol{\tau}, \mathbf{v}, q)) \to G(\lambda, \mathcal{V}) = \lambda \sum_{j=1}^{d} v_j \partial \mathbf{v}/\partial x_j - \mathbf{f}_*. \tag{5.6}$$

Clearly, the following lemma applies.

LEMMA 5.1. *$\mathcal{U} = (\boldsymbol{\sigma}, \mathbf{u}, p) \in \mathbf{X}$ is a solution of* (5.2) *if and only if $(\boldsymbol{\sigma}, \mathbf{u}, p) \in \mathbf{X}$ is a solution of* (5.4).

5.2. Div least-squares method for Navier–Stokes equations. To approximate (5.4), we introduce an approximation space $\mathbf{X}_h$ and an operator $T_h \in \mathcal{L}(\mathbf{Y}; \mathbf{X}_h)$ intended to approximate T. Set

$$F_h(\lambda, \mathcal{U}) = \mathcal{U} + T_h G(\lambda, \mathcal{U}). \tag{5.7}$$

Then, the approximation problem is to find $\mathcal{U}_h \in \mathbf{X}_h$ such that

$$F_h(\lambda, \mathcal{U}_h) = 0. \tag{5.8}$$

The following least-squares solver was introduced in [12] and we briefly present here for completeness. For the approximation space, let

$$\mathbf{X}_h = \Sigma_h^k \times V_h^{k+1} \times Q_h,$$

where Σ_h^k and V_h^{k+1} are defined in (3.7) and Q_h is defined in (4.3). To approximate solution $(\boldsymbol{\sigma}_*, \mathbf{u}_*, p_*) = T\mathbf{f}_* \in \mathbf{X}$ of (5.5), we define an approximation $(\boldsymbol{\sigma}_{*h}, \mathbf{u}_{*h}, p_{*h}) = T_h f_* \in \mathbf{X}_h$ to be the solution of the following problem:

$$\begin{aligned}(\nabla\cdot\boldsymbol{\sigma}_{*h}, \nabla\cdot\boldsymbol{\tau}_h) + (\boldsymbol{\sigma}_{*h} - \nabla\mathbf{u}_{*h} + p_{*h}, \boldsymbol{\tau}_h - \nabla\mathbf{w}_h + q_h) + (\nabla\cdot\mathbf{u}_{*h}, \nabla\cdot\mathbf{w}_h)\\ = b(\boldsymbol{\sigma}_{*h}, \mathbf{u}_{*h}, p_{*h}\,;\, \boldsymbol{\tau}_h, \mathbf{w}_h, q_h) = (\mathbf{f}_*, -\nabla\cdot\boldsymbol{\tau}_h)\end{aligned} \tag{5.9}$$

for all $(\boldsymbol{\tau}_h, \mathbf{w}_h, q_h) \in \mathbf{X}_h$. Note that the following orthogonal property holds:

$$b(\boldsymbol{\sigma}_* - \boldsymbol{\sigma}_{*h}, \mathbf{u}_* - \mathbf{u}_{*h}, p_* - p_{*h}\,;\, \boldsymbol{\tau}_h, \mathbf{w}_h, q_h) = 0 \tag{5.10}$$

for all $(\boldsymbol{\tau}_h, \mathbf{w}_h, q_h) \in \mathbf{X}_h$. Using this orthogonal property, the following error estimate is obtained. We provide a sketch of the proof.

THEOREM 5.2. *Let $(\boldsymbol{\sigma}, \mathbf{u}, p) = T\mathbf{f} \in \mathbf{X}$ be the solution of* (5.5) *and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, p_h) = T_h\mathbf{f} \in \mathbf{X}_h$ be defined in* (5.9)*. Then, the following estimate holds*

$$\begin{aligned}&\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)} + \|\mathbf{u} - \mathbf{u}_h\|_1 + \|p - p_h\|_0\\ &\qquad \le C\left(\inf_{\boldsymbol{\tau}\in\Sigma_h^k} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{H(div)} + \inf_{\mathbf{w}\in V_h^{k+1}} \|\mathbf{u} - \mathbf{w}\|_1 + \inf_{q\in Q_h} \|p - q\|_0\right).\end{aligned}$$

Proof. The following equivalence is provided in [12, Theorem 4.1]:

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(div)}^2 + \|\mathbf{u}-\mathbf{u}_h\|_1 + \|p-p_h\|_0 \simeq b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\, \mathbf{u}-\mathbf{u}_h\,, p-p_h\,;\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\, \mathbf{u}-\mathbf{u}_h\,, p-p_h).$$

Thus, using the orthogonal property (5.10), we have

$$\begin{aligned}&\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)}^2 + \|\mathbf{u} - \mathbf{u}_h\|_1 + \|p - p_h\|_0\\ &\le C\, b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h,\, \mathbf{u} - \mathbf{u}_h\,, p - p_h\,; \boldsymbol{\sigma} - \boldsymbol{\sigma}_h,\, \mathbf{u} - \mathbf{u}_h\,, p - p_h)\\ &\le C\, b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h,\, \mathbf{u} - \mathbf{u}_h\,, p - p_h\,; \boldsymbol{\sigma} - \boldsymbol{\tau},\, \mathbf{u} - \mathbf{w}\,, p - q)\\ &\le C(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)} + \|\mathbf{u} - \mathbf{u}_h\|_1 + \|p - p_h\|_0)(\|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{H(div)} + \|\mathbf{u} - \mathbf{w}\|_1 + \|p - q\|_0)\end{aligned}$$

for any $(\boldsymbol{\tau}, \mathbf{w}, q) \in \mathbf{X}_h$. This completes the proof. ∎

The following error estimate is the key for a convergence analysis for the linearization process defined in (5.4). The proof is very similar to the one given in section 4.

THEOREM 5.3. *Let* $(\boldsymbol{\sigma}, \mathbf{u}, p) = T\mathbf{f} \in \mathbf{X}$ *be the solution of* (5.5) *and* $(\boldsymbol{\sigma}_h, \mathbf{u}_h, p_h) = T_h\mathbf{f} \in \mathbf{X}_h$ *be defined in* (5.9). *Then,*

$$\|\mathbf{u} - \mathbf{u}_h\|_1 \le Ch(\|\mathbf{u}\|_2 + \|p\|_1).$$

Proof. As in the proof of Theorem 4.2, let $\mathbf{u}_I$, p_I, and $\boldsymbol{\pi}_h\boldsymbol{\sigma}$ be interpolants of $\mathbf{u}$, p, and $\boldsymbol{\sigma}$, respectively. By the triangle inequality, the approximation property of $\mathbf{u}_I$, the definition of (5.9), the coercivity, and the orthogonality property, we have

$$\begin{aligned}
\|\mathbf{u} - \mathbf{u}_h\|_1 &\le \|\mathbf{u} - \mathbf{u}_I\|_1 + \|\mathbf{u}_I - \mathbf{u}_h\|_1 \\
&\le \|\mathbf{u} - \mathbf{u}_I\|_1 + \|\mathbf{u}_I - \mathbf{u}_h\|_1 + \|\boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|p_I - p_h\|_0 \\
&\le Ch\|\mathbf{u}\|_2 + \|\mathbf{u}_I - \mathbf{u}_h\|_1 + \|\boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|p_I - p_h\|_0 \\
&= Ch\|\mathbf{u}\|_2 + b(\boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h, p_I - p_h\,;\, \boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h, p_I - p_h)^{\frac{1}{2}} \\
&= Ch\|\mathbf{u}\|_2 + b(\boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}, \mathbf{u}_I - \mathbf{u}, p_I - p\,;\, \boldsymbol{\pi}_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u}_I - \mathbf{u}_h, p_I - p_h)^{\frac{1}{2}}.
\end{aligned}$$

The rest of the proof is almost identical to that of Theorem 4.2. ◻

Now, suppose that (5.4) has a branch of nonsingular solution $\{(\lambda, \mathcal{U}(\lambda)) \,:\, \lambda \in \Lambda\}$. We need to make additional assumptions in order to use the abstract approximation theory of [19]. First, we suppose that there exists another Banach space $\mathbf{Z}$ contained in $\mathbf{Y}$, with continuous imbedding, such that

$$D_{\mathcal{U}}G(\lambda, \mathcal{U}) \in \mathcal{L}(\mathbf{X}; \mathbf{Z}) \quad \forall\, \lambda \in \Lambda, \quad \forall \mathcal{U} \in \mathbf{X}. \tag{5.11}$$

Next, concerning the approximation properties of the operator T_h, we assume that

$$\lim_{h \to \infty} \|(T - T_h)g\|_{\mathbf{X}} = 0 \quad \forall\, g \in \mathbf{Y} \tag{5.12}$$

and

$$\lim_{h \to \infty} \|T - T_h\|_{\mathcal{L}(\mathbf{Z};\mathbf{X})} = 0. \tag{5.13}$$

We will verify these assumptions later. Under these assumptions, the following theorem is proved in [19, Theorem 3.3, pp. 307].

THEOREM 5.4. *Assume that* G *is a* $\mathcal{C}^2$*-mapping from* $\Lambda \times \mathbf{X}$ *into* $\mathbf{Y}$ *and the mapping* D^2G *is bounded on all bounded subsets of* $\Lambda \times \mathbf{X}$. *Assume, in addition, that the conditions* (5.11), (5.12), *and* (5.13) *hold and that* $\{(\lambda, \mathcal{U}(\lambda)) \,:\, \lambda \in \Lambda\}$ *is a branch of nonsingular solutions of* (5.4). *Then for* $h \le h_0$ *small enough, there exists a neighborhood* $\mathcal{O}$ *of the origin in* $\mathbf{X}$ *and a unique* $\mathcal{C}^2$*-function* $\lambda \to \mathcal{U}_h(\lambda) \in \mathbf{X}$ *such that* $\{(\lambda, \mathcal{U}_h(\lambda)) : \lambda \in \Lambda\}$ *is a branch of nonsingular solutions of* (5.7) *and that*

$$\mathcal{U}_h(\lambda) - \mathcal{U}(\lambda) \in \mathcal{O} \quad \forall\, \lambda \in \Lambda.$$

Furthermore, there exists a constant $K > 0$ *independent of* h *and* λ *with*

$$\|\mathcal{U}_h(\lambda) - \mathcal{U}(\lambda)\|_{\mathbf{X}} \le K\|(T_h - T)G(\lambda, \mathcal{U}(\lambda))\|_{\mathbf{X}} \quad \forall \lambda \in \Lambda.$$

In order to apply Theorem 5.4, we need to verify the assumptions made in the statement, and we provide the proof here.

LEMMA 5.5. *Equation* (5.4) *has a branch of nonsingular solutions,* $\{(\lambda, \mathcal{U}(\lambda)) : \lambda \in \Lambda\} \in \mathbf{X}$.

Proof. For all $\mathcal{V} = (\boldsymbol{\tau}, \mathbf{v}, q) \in \mathbf{X}$, the Fréchet derivative of G with respect to $\mathcal{U}$ is

$$D_{\mathcal{U}} G(\lambda, \mathcal{U}) \cdot \mathcal{V} = \lambda \sum_{j=1}^{d} (u_j \partial \mathbf{v} / \partial x_j + v_j \partial \mathbf{u} / \partial x_j).$$

Hence, $\mathcal{U} = (\boldsymbol{\sigma}, \mathbf{u}, p)$ is a nonsingular solution of (5.4), or equivalently $D_{\mathcal{U}} F(1/\nu, \mathbf{U})$ is an isomorphism of $\mathbf{X}$ if and only if, for each $\mathcal{W} = (\boldsymbol{\psi}, \mathbf{w}, r) \in \mathbf{X}$, there exists a unique $\mathcal{V} = (\boldsymbol{\tau}, \mathbf{v}, q) \in \mathbf{X}$ such that

$$(5.14) \qquad \begin{cases} -\nabla \cdot (\boldsymbol{\tau} - \boldsymbol{\psi}) + \frac{1}{\nu} \sum_{j=1}^{d} (u_j \partial \mathbf{v} / \partial x_j + v_j \partial \mathbf{u} / \partial x_j) = \mathbf{0} & \text{in } \Omega, \\ (\boldsymbol{\tau} - \boldsymbol{\psi}) - \nabla (\mathbf{v} - \mathbf{w}) + (q - r)\, I = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot (\mathbf{v} - \mathbf{w}) = 0 & \text{in } \Omega, \\ \mathbf{v} - \mathbf{w} = \mathbf{0} & \text{on } \partial\Omega. \end{cases}$$

Eliminating $\boldsymbol{\tau} - \boldsymbol{\psi}$ in the above equations and letting $\mathbf{s} = \mathbf{v} - \mathbf{w}$ and $\bar{q} = q - r$, then (5.14) may be rewritten as

$$(5.15) \qquad \begin{cases} -\triangle \mathbf{s} + \nabla \bar{q} + \dfrac{1}{\nu} \displaystyle\sum_{j=1}^{d} (u_j \partial \mathbf{s} / \partial x_j + s_j \partial \mathbf{u} / \partial x_j) = \widetilde{\mathbf{g}} & \text{in } \Omega, \\ \nabla \cdot \mathbf{s} = 0 & \text{in } \Omega, \\ \mathbf{s} = 0 & \text{on } \partial\Omega, \end{cases}$$

where $\widetilde{\mathbf{g}} = -\sum_{j=1}^{d} (u_j \partial \mathbf{w} / \partial x_j + w_j \partial \mathbf{u} / \partial x_j) \in L^2(\Omega)$. The uniqueness and existence of the solution $(\mathbf{s}, \bar{q}) \in W_2^1 \times L_0^2$ is well-known [19, page 300, Lemma 3.2]. Now, the H^2 regularity assumption, $(\mathbf{s}, \bar{q}) \in W_2^2 \times W_2^1 \subset W_2^{\frac{3}{2}-\delta} \times L^2$, implies that $\boldsymbol{\tau} - \boldsymbol{\psi} \in H(div)^d$. This, in turn, yields that (5.14) has a unique solution $\mathcal{V} = (\boldsymbol{\tau}, \mathbf{v}, q) \in \mathbf{X}$ and, hence, (5.4) has a branch of nonsingular solutions. ∎

To prove assumption (5.11), take

$$(5.16) \qquad \mathbf{Z} = W_2^{\epsilon},$$

where $\epsilon > 0$ is sufficiently small.

LEMMA 5.6. *For* $\mathcal{U} = (\boldsymbol{\sigma}, \mathbf{u}, p) \in \mathbf{X} = H(div) \times W_2^{\frac{3}{2}-\delta} \times L_0^2$,

$$D_{\mathcal{U}} G(\lambda, \mathcal{U}) \in \mathcal{L}(\mathbf{X}, \mathbf{Z}).$$

Proof. Note that for any $\mathcal{V} = (\boldsymbol{\tau}, \mathbf{v}, q) \in \mathbf{X}$, we have

$$D_{\mathcal{U}} G(\lambda, \mathcal{U}) \cdot \mathcal{V} = \lambda \sum_{j=1}^{d} (u_j \partial \mathbf{v} / \partial x_j + v_j \partial \mathbf{u} / \partial x_j).$$

Thus, it is necessary to show that

$$\|u_j \partial \mathbf{v} / \partial x_j\|_{W_2^{\epsilon}},\ \|v_j \partial \mathbf{u} / \partial x_j\|_{W_2^{\epsilon}} \le C(\|\boldsymbol{\tau}\|_{H(div)} + \|\mathbf{v}\|_{W_2^{\frac{3}{2}-\delta}} + \|q\|_0),$$

where C is independent of $\mathcal{V}$ but may depend on $\mathcal{U}$. Obviously, it suffices to establish that

$$\|u_j \partial \mathbf{v}/\partial x_j\|_{W_2^\epsilon} \le C(u_j)\,\|\mathbf{v}\|_{W_2^{\frac{3}{2}-\delta}} \quad \text{and} \quad \|v_j \partial \mathbf{u}/\partial x_j\|_{W_2^\epsilon} \le C(\mathbf{u})\,\|\mathbf{v}\|_{W_2^{\frac{3}{2}-\delta}}$$

for sufficiently small ϵ and δ. For $d = 2$, the above inequalities can be easily obtained by taking L^∞ norm for u_j or v_j and applying Sobolev's inequality. For $d = 3$, we start with the first inequality. Using Hölder's inequality, we have

$$\|u_j \partial \mathbf{v}/\partial x_j\|_{W_2^\epsilon} \le \|u_j\|_{W_{2p}^\epsilon} \|\mathbf{v}\|_{W_{2q}^{1+\epsilon}} \quad \text{with } \frac{1}{p} = 1 - \frac{1}{q}.$$

By taking $q = \frac{3}{2(1+\delta+\epsilon)}$ in the Sobolev imbedding theorem, we have

$$\|\mathbf{v}\|_{W_{2q}^{1+\epsilon}} \le C\,\|\mathbf{v}\|_{W_2^{3/2-\delta}}.$$

By the choice of q, we have $p = \frac{3}{1-2(\delta+\epsilon)}$. Again, by the Sobolev imbedding theorem, we have

$$W_2^{\frac{3}{2}-\delta} \subset W^\epsilon_{\frac{3}{\delta+\epsilon}} \subset W_{2p}^\epsilon$$

for δ and ϵ being sufficiently small. That is, $\|u_j\|_{W_{2p}^\epsilon}$ is finite since $\mathbf{u} \in W_2^{\frac{3}{2}-\delta}$. This proves

$$\|v_j \partial \mathbf{u}/\partial x_j\|_{W_2^\epsilon} \le C(\mathbf{u})\|\mathbf{v}\|_{W_2^{3/2-\delta}}.$$

For the second inequality, again using Hölder's inequality, we have

$$\|v_j \partial \mathbf{u}/\partial x_j\|_{W_2^\epsilon} \le \|v_j\|_{W_{2p}^\epsilon} \|\mathbf{u}\|_{W_{2q}^{1+\epsilon}},$$

where $\frac{1}{p} = 1 - \frac{1}{q}$. By taking $q = \frac{3}{2(1+\delta+\epsilon)}$ in the Sobolev imbedding theorem, we have

$$\|\mathbf{u}\|_{W_{2q}^{1+\epsilon}} \le C\|\mathbf{u}\|_{W_2^{3/2-\delta}}.$$

By the choice of q, we have $p = \frac{3}{1-2(\delta+\epsilon)}$. By Sobolev imbedding theorem, we have

$$W_2^{\frac{3}{2}-\delta} \subset W^\epsilon_{\frac{3}{\epsilon+\delta}} \subset W^\epsilon_{\frac{6}{1-2(\delta+\epsilon)}} = W_{2p}^\epsilon$$

for δ and ϵ being sufficiently small. That is, $\|v_j\|_{W_{2p}^\epsilon} \le C\|\mathbf{v}\|_{W_2^{\frac{3}{2}-\delta}}$. This proves

$$\|v_j \partial \mathbf{u}/\partial x_j\|_{W_2^\epsilon} \le C(\mathbf{u})\|\mathbf{v}\|_{W_2^{3/2-\delta}}$$

and, hence, completes the proof of the lemma. □

By the Sobolev imbedding theorem, $\mathbf{Z} = W_2^\epsilon$ for all $\epsilon > 0$ is compactly imbedded into $\mathbf{Y} = L^2$. Thus, (5.13) follows if we prove (5.12).

LEMMA 5.7. *For any $g \in \mathbf{Y} = L^2$,*

$$\lim_{h\to\infty} \|(T - T_h)g\|_{\mathbf{X}} = 0.$$

Proof. Let $\mathcal{U} = (\boldsymbol{\sigma}, \mathbf{u}, p) = Tg$ and $\mathcal{U}_h = (\boldsymbol{\sigma}_h, \mathbf{u}_h, p_h) = T_h g$. Let $\mathbf{u}_I \in V_h^1$ be the interpolant of $\mathbf{u}$. By the triangle inequality, the approximation property, the inverse inequality, and Theorem 5.3, we have

$$\begin{aligned}
\|\mathbf{u} - \mathbf{u}_h\|_{W_2^{\frac{3}{2}-\delta}} &\le \|\mathbf{u} - \mathbf{u}_I\|_{W_2^{\frac{3}{2}-\delta}} + \|\mathbf{u}_I - \mathbf{u}_h\|_{W_2^{\frac{3}{2}-\delta}} \\
&\le Ch^{\frac{1}{2}+\delta}\|\mathbf{u}\|_2 + h^{-\frac{1}{2}+\delta}\|\mathbf{u}_I - \mathbf{u}_h\|_{W_2^1} \\
&\le Ch^{\frac{1}{2}+\delta}\|\mathbf{u}\|_2 + h^{-\frac{1}{2}+\delta}(\|\mathbf{u}_I - \mathbf{u}\|_{W_2^1} + \|\mathbf{u} - \mathbf{u}_h\|_{W_2^1}) \\
&\le Ch^{\frac{1}{2}+\delta}\|\mathbf{u}\|_2 + h^{-\frac{1}{2}+\delta}\|\mathbf{u} - \mathbf{u}_h\|_{W_2^1} \le Ch^{\frac{1}{2}+\delta}(\|\mathbf{u}\|_2 + \|p\|_1).
\end{aligned}$$

Hence, it follows from the definition of $\mathbf{X}$, the above inequality, and Theorem 5.2 that

$$\begin{aligned}
\|\mathcal{U} - \mathcal{U}_h\|_{\mathbf{X}} &\le C(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)} + \|\mathbf{u} - \mathbf{u}_h\|_{W_2^{\frac{3}{2}-\delta}} + \|p - p_h\|_0) \\
&\le C(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(div)} + \|p - p_h\|_0) + Ch^{\frac{1}{2}+\delta}(\|\mathbf{u}\|_2 + \|p\|_1) \\
&\le C\left(\inf_{\boldsymbol{\tau}\in\Sigma_h^k}\|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{H(div)} + \inf_{\mathbf{w}\in V_h^{k+1}}\|\mathbf{u} - \mathbf{w}\|_1 + \inf_{q\in Q_h}\|p - q\|_0\right) \\
&\quad + Ch^{\frac{1}{2}+\delta}(\|\mathbf{u}\|_2 + \|p\|_1),
\end{aligned}$$

which implies that $\|\mathcal{U} - \mathcal{U}_h\|_{\mathbf{X}} \to 0$ as $h \to 0$. This completes the proof of the lemma. ◻

Remark 5.8. Under the same assumption as in Theorem 5.4 and taking $g = G(\lambda, \mathcal{U}(\lambda))$ in Lemma 5.7, we have the following error estimates:

$$\begin{aligned}
\|\mathcal{U}_h(\lambda) - \mathcal{U}(\lambda)\|_{\mathbf{X}} \le C\Bigg(&\inf_{\boldsymbol{\tau}\in\Sigma_h^k}\|\boldsymbol{\sigma}(\lambda) - \boldsymbol{\tau}\|_{H(div)} + \inf_{\mathbf{w}\in V_h^{k+1}}\|\mathbf{u}(\lambda) - \mathbf{w}\|_1 \\
&+ \inf_{q\in Q_h}\|p(\lambda) - q\|_0\Bigg) + Ch^{\frac{1}{2}+\delta}(\|\mathbf{u}(\lambda)\|_2 + \|p(\lambda)\|_1).
\end{aligned}$$

REFERENCES

[1] A. Aziz, R. Kellogg, and A. Stephens, *Least-squares methods for elliptic systems*, Math. Comp., 44 (1985), pp. 53–70.

[2] P. B. Bochev, *Analysis of least-squares finite element methods for the Navier–Stokes equations*, SIAM J. Numer. Anal., 34 (1997), pp. 1817–1844.

[3] P. B. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux first-order system least-squares principles for the Navier–Stokes equations; Part* I, SIAM J. Numer. Anal., 35 (1998), pp. 990–1009.

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

[5] P. B. Bochev and M. D. Gunzburger, *On least-squares finite element methods for the Poisson equation and their connection to the Dirichlet and Kelvin principles*, SIAM J. Numer. Anal., 43 (2005), pp. 340–362.

[6] P. B. Bochev and M. D. Gunzburger, *A locally conservative least-squares method for Darcy flows*, Comm. Numer. Methods Engrg., 24 (2008), pp. 97–110.

[7] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order system*, Math. Comp., 66 (1997), pp. 935–955.

[8] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Spring-Verlag, Berlin, 1991.

[9] Z. Cai and J. Ku, *The L^2 norm error estimates for the div least-squares method*, SIAM J. Numer. Anal., 44 (2006), pp. 1721–1734.

[10] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part* I, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.

[11] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part* II, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[12] Z. Cai, B. Lee, and P. Wang, *Least-squares methods for incompressible Newtonian fluid flow: Linear stationary problems*, SIAM J. Numer. Anal., 42 (2004), pp. 843–859.

[13] Z. Cai and G. Starke, *Least-squares methods for linear elasticity*, SIAM J. Numer. Anal., 42 (2004), pp. 826–842.

[14] G. F. Carey and Y. Shen, *Convergence studies of least-squares finite elements for first order systems*, Comm. Appl. Numer. Methods, 5 (1989), pp. 427–434.

[15] C. L. Chang, *A least-squares finite element method for the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 83 (1990), pp. 1–7.

[16] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[17] C. L. Cox and G. Fix, *On the accuracy of least squares methods in the presence of corner singularities*, Comput. Math. Appl., 10 (1984), pp. 463–475.

[18] G. Fix, M. Gunzburger, and R. Nicolaides, *On the finite element methods of least-squares type*, Comput. Math. Appl., 5 (1979), pp. 87–98.

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

[20] D. C. Jespersen, *A least-square decomposition method for solving elliptic systems*, Math. Comp., 31 (1977), pp. 873–880.

[21] B. N. Jiang and L. A. Povinelli, *Optimal least-squares finite element method for elliptic problems*, Comput. Meth. Appl. Mech. Engrg., 102 (1993), pp. 199–212.

[22] J. Ku, *Weak coupling of solutions of least-squares method*, Math. Comp., 77 (2008), pp. 1323–1332.

[23] P. P. Lynn and S. K. Arya, *Use of the least squares criterion in the finite element formulation*, Internat. J. Numer. Methods Engrg., 6 (1973), pp. 75–88.

[24] T. Manteuffel, S. McCormick, and C. Pflaum, *Improved discretization error estimates for first-order system least squares*, J. Numer. Math. 11 (2003), pp. 163–177.

[25] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1368–1377.

[26] A. I. Pehlivanov, G. F. Carey, R. D. Lazarov, and Y. Shen, *Convergence of least squares finite elements for first order ODE systems*, Computing, 51 (1993), pp. 111–123.

[27] A. I. Pehlivanov, G. F. Carey, and P. S. Vassilievski, *Least-squares mixed finite element methods for non-selfadjoint elliptic problems:* I. *Error extimates*, Numer. Math. 72 (1996), pp. 501–522.

[28] D. Yang, *Analysis of least-squares mixed finite element methods for nonlinear nonstationary convection-diffusion problems*, Math. Comp., 69 (2000), pp. 929–963.