

First-Order System Least-Squares Methods for a Flux Control Problem by the Stokes Flow

Soorok Ryu¹, Sang Dong Kim² and Hyung-Chun Lee^{3,*}

¹ Department of Industrial and Applied Mathematics, Kyungpook National University, Daegu 702-701, Korea.

² Department of Mathematics, Kyungpook National University, Daegu 702-701, Korea.

³ Department of Mathematics, Ajou University, Suwon 443-749, Korea.

Received 23 April 2009; Accepted (in revised version) 4 September 2009

Available online 28 October 2009

Abstract. This article deals with a first-order least-squares approach to the solution of an optimal control problem governed by Stokes equations. As with our earlier work on a velocity control by the Stokes flow in [S. Ryu, H.-C. Lee and S. D. Kim, *SIAM J. Numer. Anal.*, 47 (2009), pp. 1524-1545], we recast the objective functional as a H^1 seminorm in the velocity control term. By introducing a *velocity-flux* variable and using the Lagrange multiplier rule, a first-order optimality system is obtained. We show that the least-squares principle based on L^2 norms applied to this system yields the optimal discretization error estimates for each variable in H^1 norm, including the velocity flux. For numerical tests, multigrid method is employed to the discrete algebraic system, so that the velocity and flux controls are obtained.

AMS subject classifications: 65M55, 65N30, 49J20, 49K20

Key words: Optimal control, optimization, least-squares finite element methods, Lagrange multiplier, Stokes equations.

1 Introduction

Optimal control problems governed by partial differential equations (PDEs) can be reduced to a system of coupled PDEs by the Lagrange multiplier method [13, 14, 18, 25]. Such system of coupled PDEs and optimal control problems have been interesting subjects because of not only their importance in the design process but also the needs of efficient and numerical methods for implementations.

*Corresponding author. *Email addresses:* sryu@knu.ac.kr (S. Ryu), skim@knu.ac.kr (S. D. Kim), hclee@ajou.ac.kr (H.-C. Lee)

There have been considerable attentions in first-order system least-squares (FOSLS) approaches for Navier-Stokes or Stokes equations in many literatures [6, 9, 10, 12, 20–23]. These principles result in symmetric positive definite algebraic systems. Moreover, they enable us to avoid using the finite elements satisfying the inf-sup condition. Some applications of least-squares finite element methods to optimization problem have been previously discussed in [3–5, 7, 8, 15, 24, 27]. In particular, [8] developed an abstract form of the FOSLS for optimal control problems governed by elliptic PDEs. Recently, [27] provides a nice synthesis of optimal control by Stokes flows and FOSLS principles by drawing on [8] and [12]. In [27], they only considered a velocity control using L^2 norms in the objective functional. In this work, as an improved version of [27], we consider a flux control using the cost functional which has the H^1 seminorm as a velocity control term.

From the Poincaré-Friedrichs inequality, the velocity control is also obtained by a flux control. In a different way with the one in [27], the Lagrange multiplier method is used after introducing a new variable $\mathbf{U} = \nabla \mathbf{u}^t$, so that the value \mathbf{V} is a Lagrange multiplier and $\mathbf{V} \neq \nabla \mathbf{v}^t$.

The objective functional we consider is

$$\mathcal{J}(\mathbf{u}, \mathbf{f}) = \frac{1}{2} \int_{\Omega} |\mathbf{u} - \hat{\mathbf{u}}|^2 + |\nabla \mathbf{u}^t - \nabla \hat{\mathbf{u}}^t|^2 dx + \frac{\sigma}{2} \int_{\Omega} |\mathbf{f}|^2 dx, \quad (1.1)$$

where $(\cdot)^t$ denotes the transpose, $\hat{\mathbf{u}}$ is the given target velocity, σ is a positive penalty parameter. The constraint is the Stokes equations such that

$$\begin{cases} -\nu \Delta \mathbf{u} + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on } \partial\Omega, \\ \int_{\Omega} p dx = 0, \end{cases} \quad (1.2)$$

where \mathbf{u} and p denote the velocity and pressure, respectively, ν the constant kinematic viscosity, and \mathbf{f} the control function. Here $\Omega \subset \mathbb{R}^n$ ($n = 2$ or 3) is a bounded convex polyhedron or has $C^{1,1}$ boundary. The problem we study is to find an optimal state (\mathbf{u}, p) and an optimal control \mathbf{f} which minimize the H^1 -norm distance between \mathbf{u} and $\hat{\mathbf{u}}$ satisfying the Stokes system (1.2). If Ω is connected and bounded at least in one direction and $\hat{\mathbf{u}} \in [H_0^1(\Omega)]^n$, then there exists a constant $C(\Omega)$ such that

$$\|\mathbf{u} - \hat{\mathbf{u}}\| < C(\Omega) \|\nabla \mathbf{u}^t - \nabla \hat{\mathbf{u}}^t\|, \quad (1.3)$$

by the Poincaré-Friedrichs inequality (see [17]). From (1.3), the objective functional (1.1) can be replaced by

$$\mathcal{J}(\mathbf{u}, \mathbf{f}) = \frac{1}{2} \int_{\Omega} |\nabla \mathbf{u}^t - \nabla \hat{\mathbf{u}}^t|^2 dx + \frac{\sigma}{2} \int_{\Omega} |\mathbf{f}|^2 dx. \quad (1.4)$$

Hence, we consider the objective functional (1.4) instead of (1.1). Such a constrained optimization problem can be converted to an unconstrained optimization problem by the Lagrange multiplier rule [8, 13]. Before using the Lagrange multiplier rule, we take the techniques in [12] by introducing the flux variable $\mathbf{U} = \nabla \mathbf{u}^t$ so that the Stokes system is a first-order system. Then we can get the optimality system which consists of the *state system*, an *adjoint system* and an *optimality condition* using the Lagrange multiplier rule. This optimality system is a coupled first-order system.

We first consider the H^{-1} - L^2 norm based least-squares functional corresponding to the first-order coupled system. Using the norm equivalence of this functional, we show the H^1 norm equivalence of the L^2 based functional (see Section 3). Note that we impose the same weight as the one in [27] on the least-squares functional to get desired numerical results.

This paper is organized as follows. In Section 2, we present the first-order optimality system from the given constrained optimal control problem using the Lagrange multiplier rule; in Section 3, two types of first-order systems least-squares functionals are defined and then the existence and uniqueness of the solutions are shown. In Section 4, optimal discretization error estimates in the H^1 norm are proved. Some numerical results are presented for a simple example in Section 5. Finally, concluding remarks are added in Section 6.

2 Coupled first-order system formulations

In this paper, the standard Sobolev space notation is used; $H^s(\Omega)$ with their associated inner products $(\cdot, \cdot)_s$ and norms $\|\cdot\|_s, s \geq 0$ will be used. For example, $H^0(\Omega)$ is the usual $L^2(\Omega)$ with the norm $\|\cdot\|_0 = \|\cdot\|$ and inner product (\cdot, \cdot) . The $L_0^2(\Omega)$ is the subspace of all square integrable functions with zero mean. The space $H^{-1}(\Omega)$ denotes the dual of $H_0^1(\Omega)$ equipped with norm $\|\phi\|_{-1} = \sup_{0 \neq v \in H_0^1(\Omega)} (\phi, v) / \|v\|_1$. Finally, the standard *div* space $H(\text{div}; \Omega)$ and *curl* space $H(\mathbf{curl}; \Omega)$ will be used (see [1, 17]).

For first-order formulations, we use velocity-velocity gradient-pressure formulation in [12] by introducing the state velocity flux variable $\mathbf{U} = \nabla \mathbf{u}^t = (\nabla u_1, \dots, \nabla u_n)$. Then the system (1.2) and (1.4) may be the following minimization problem: minimize the objective functional

$$\mathcal{J}(\mathbf{U}, \mathbf{f}) = \frac{1}{2} \|\mathbf{U} - \nabla \hat{\mathbf{u}}^t\|_1^2 + \frac{\sigma}{2} \|\mathbf{f}\|^2, \quad (2.1)$$

subject to

$$\left\{ \begin{array}{ll} -v(\nabla \cdot \mathbf{U})^t + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \mathbf{U} - \nabla \mathbf{u}^t = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on } \partial\Omega, \\ \int_{\Omega} p \, dx = 0, & \end{array} \right. \quad (2.2)$$

where $|\cdot|_1$ denotes the $H^1(\Omega)$ seminorm.

By introducing the Lagrange multipliers \mathbf{v}, \mathbf{V} and q , the constrained optimization problem (2.1) and (2.2) may be converted as the following equivalent coupled system (see [13]):

$$\left\{ \begin{array}{ll} -\nu(\nabla \cdot \mathbf{U})^t + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \mathbf{U} - \nabla \mathbf{u}^t = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on } \partial\Omega, \\ (\nabla \cdot \mathbf{V})^t + \nabla q = \mathbf{0} & \text{in } \Omega, \\ \mathbf{V} - \nu \nabla \mathbf{v}^t + \mathbf{U} = \nabla \hat{\mathbf{u}}^t & \text{in } \Omega, \\ \nabla \cdot \mathbf{v} = 0 & \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{on } \partial\Omega, \end{array} \right. \quad (2.3)$$

and

$$-\frac{1}{\sigma} \mathbf{v} = \mathbf{f}. \quad (2.4)$$

We can get rid of the control \mathbf{f} from the state system in (2.3) by the relation of the control and the adjoint variable in (2.4). Following the results in [12], we obtain the extended optimality system for (2.3) and (2.4):

$$\left\{ \begin{array}{ll} -\nu(\nabla \cdot \mathbf{U})^t + \nabla p + \frac{\mathbf{v}}{\sigma} = \mathbf{0} & \text{in } \Omega, \\ \mathbf{U} - \nabla \mathbf{u}^t = \mathbf{0} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \nabla \times \mathbf{U} = \mathbf{0} & \text{in } \Omega, \\ \nabla(\text{tr} \mathbf{U}) = 0 & \text{in } \Omega, \\ (\nabla \cdot \mathbf{V})^t + \nabla q = \mathbf{0} & \text{in } \Omega, \\ \mathbf{V} - \nu \nabla \mathbf{v}^t + \mathbf{U} = \nabla \hat{\mathbf{u}}^t & \text{in } \Omega, \\ \nabla \cdot \mathbf{v} = 0 & \text{in } \Omega, \\ \nabla \times \mathbf{V} = \mathbf{0} & \text{in } \Omega, \\ \nabla(\text{tr} \mathbf{V}) = 0 & \text{in } \Omega, \end{array} \right. \quad (2.5)$$

with $\mathbf{u} = \mathbf{v} = \mathbf{0}, \mathbf{n} \times \mathbf{U} = \mathbf{n} \times \mathbf{V} = \mathbf{0}$ on $\partial\Omega$ and

$$\int_{\Omega} p \, dx = \int_{\Omega} q \, dx = 0.$$

Remark 2.1. As a more general case, one can regard the nonhomogeneous velocity boundary condition problem: find $(\mathbf{w}, p, \mathbf{f}) \in [H^1(\Omega)]^n \times L_0^2(\Omega) \times [L^2(\Omega)]^n$ minimizing the functional

$$\mathcal{J}(\mathbf{w}, \mathbf{f}) = \frac{1}{2} \|\mathbf{w} - \hat{\mathbf{w}}\|_1^2 + \frac{\sigma}{2} \|\mathbf{f}\|, \quad (2.6)$$

subject to

$$\begin{cases} -\nu\Delta\mathbf{w} + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{w} = 0 & \text{in } \Omega, \\ \mathbf{w} = \mathbf{g} & \text{on } \Gamma := \partial\Omega, \end{cases} \quad (2.7)$$

where $\mathbf{g} \in [H^{1/2}(\Gamma)]^n$ such that $\int_{\Gamma} \mathbf{g} \cdot \mathbf{n} \, ds = 0$. It is well known that if \mathbf{w} is the solution of Stokes equations (2.7) then it can be expressed as $\mathbf{w} = \mathbf{u}_0 + \mathbf{u}$ (see [17]), where $\mathbf{u} \in [H_0^1(\Omega)]^n$ is the velocity solution of (1.2) with $\mathbf{f} + \nu\Delta\mathbf{u}_0$ in a forcing term and $\mathbf{u}_0 \in [H^1(\Omega)]^n$ is a function such that

$$\nabla \cdot \mathbf{u}_0 = 0 \quad \text{in } \Omega, \quad \mathbf{u}_0 = \mathbf{g} \quad \text{on } \Gamma.$$

Then, the minimizing objective functional and the state system will be

$$\mathcal{J}(\mathbf{u}, \mathbf{f}) = \frac{1}{2} \|\mathbf{u} - \hat{\mathbf{u}}\|_1^2 + \frac{\sigma}{2} \|\mathbf{f}\|,$$

and

$$\begin{cases} -\nu\Delta\mathbf{u} + \nabla p = \mathbf{f} + \nu\Delta\mathbf{u}_0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on } \Gamma, \end{cases}$$

where, $\hat{\mathbf{u}} = \hat{\mathbf{w}} - \mathbf{u}_0$. Similarly, by introducing a velocity flux variable $\mathbf{U} = \nabla\mathbf{u}^t$ and the Lagrange multiplier method, one can induce the first optimality system. In this case, one may have

$$-\nu(\nabla \cdot \mathbf{U})^t + \nabla p + \frac{\mathbf{v}}{\sigma} = \nu\Delta\mathbf{u}_0$$

and

$$(\nabla \cdot \mathbf{V})^t + \nabla q + \mathbf{u} = \hat{\mathbf{u}}$$

instead of the first line and the sixth line in Eq. (2.5), respectively. Since the latter equation is the same one in [27], one can easily modified the above case using the results in this paper and [27]. If $\hat{\mathbf{u}} := \hat{\mathbf{w}} - \mathbf{u}_0 \in [H_0^1(\Omega)]^2$, then only the first line in (2.5) is changed into

$$-\nu(\nabla \cdot \mathbf{U})^t + \nabla p + \frac{\mathbf{v}}{\sigma} = \nu\Delta\mathbf{u}_0.$$

In this sense, it is enough to consider a homogeneous velocity boundary condition for our optimal control problem.

3 Norm equivalences of first-order least-squares functionals

In this section we consider two types of least-squares functionals using different norms, one of which employs H^{-1} - L^2 norms based on the system (2.3) and the other of which employs only L^2 norms for (2.5). The weight ν^2 will be given for the residuals of state flux variable \mathbf{U} and $\nabla\mathbf{u}^t$ and for the residual for the state continuity equation $\nabla \cdot \mathbf{u}$ as

suggested in [12]. Accordingly, the weight $\frac{1}{\sigma^2}$ is used to adjoint system for balancing the weights between the residuals from the state and adjoint systems (see [27]). In this sense, the first least-squares functional is defined for the system (2.3) as

$$\begin{aligned} & \mathcal{L}_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \hat{\mathbf{u}}) \\ &= \left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \right\|_{-1}^2 + \nu^2 \|\mathbf{U} - \nabla \mathbf{u}^t\|^2 + \nu^2 \|\nabla \cdot \mathbf{u}\|^2 \\ & \quad + \frac{\nu^2}{\sigma^2} \|(\nabla \cdot \mathbf{V})^t + \nabla q\|_{-1}^2 + \frac{1}{\sigma^2} \|\mathbf{V} - \nu \nabla \mathbf{v}^t + \mathbf{U} - \nabla \hat{\mathbf{u}}^t\|^2 + \frac{\nu^2}{\sigma^2} \|\nabla \cdot \mathbf{v}\|^2, \end{aligned}$$

and the second functional for the extended system (2.5) is defined as

$$\begin{aligned} & \mathcal{L}_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \hat{\mathbf{u}}) \\ &= \left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \right\|_{-1}^2 + \nu^2 \|\mathbf{U} - \nabla \mathbf{u}^t\|^2 + \nu^2 \|\nabla \cdot \mathbf{u}\|^2 + \nu^2 \|\nabla \times \mathbf{U}\|^2 \\ & \quad + \nu^2 \|\nabla \text{tr}(\mathbf{U})\|^2 + \frac{\nu^2}{\sigma^2} \|(\nabla \cdot \mathbf{V})^t + \nabla q\|^2 + \frac{1}{\sigma^2} \|\mathbf{V} - \nu \nabla \mathbf{v}^t + \mathbf{U} - \nabla \hat{\mathbf{u}}^t\|^2 \\ & \quad + \frac{\nu^2}{\sigma^2} \|\nabla \cdot \mathbf{v}\|^2 + \frac{\nu^2}{\sigma^2} \|\nabla \times \mathbf{V}\|^2 + \frac{\nu^2}{\sigma^2} \|\nabla \text{tr}(\mathbf{V})\|^2. \end{aligned}$$

For the proof of a product H^1 norm equivalence of the \mathcal{L}_2 functional, we first show the coercivity of the \mathcal{L}_1 functional in a product L^2 - H^1 norm as a vehicle. Then we show the existence of solution minimizing the quadratic functionals $\mathcal{L}_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \hat{\mathbf{u}})$ over the proper solution spaces $\mathcal{W} \times \mathcal{W}$, which will be defined later. In other words, we want to find $(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \in \mathcal{W} \times \mathcal{W}$ such that

$$(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) = \arg \inf_{(\tilde{\mathbf{U}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\mathbf{V}}, \tilde{\mathbf{v}}, \tilde{q}; \hat{\mathbf{u}}) \in \mathcal{W} \times \mathcal{W}} \mathcal{L}_2(\tilde{\mathbf{U}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\mathbf{V}}, \tilde{\mathbf{v}}, \tilde{q}; \hat{\mathbf{u}}). \tag{3.1}$$

Let

$$\mathbf{W}_1 = H(\text{div}; \Omega)^n \times H_0^1(\Omega)^n \times (L_0^2(\Omega) \cap H^1(\Omega)).$$

Lemma 3.1. *For any $(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \in \mathbf{W}_1 \times \mathbf{W}_1$, there exists a positive constant C such that*

$$\|p\| \leq C \left(\left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \right\|_{-1} + \nu \|\mathbf{U}\| + \frac{1}{\sigma} \|\mathbf{v}\|_{-1} \right), \tag{3.2}$$

and

$$\|q\| \leq C \left(\|(\nabla \cdot \mathbf{V})^t + \nabla q\|_{-1} + \|\mathbf{V}\| \right). \tag{3.3}$$

Proof. Since one may modify the proof of Lemma 3.1 in [27], we omit the proof. □

For convenience, we denote M_1 and M_2 as the following norms:

$$M_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) = \|\mathbf{U}\|^2 + \|\mathbf{u}\|_1^2 + \|p\|^2 + \|\mathbf{V}\|^2 + \|\mathbf{v}\|_1^2 + \|q\|^2,$$

and

$$M_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) = \|\mathbf{U}\|_1^2 + \|\mathbf{u}\|_1^2 + \|p\|_1^2 + \|\mathbf{V}\|_1^2 + \|\mathbf{v}\|_1^2 + \|q\|_1^2.$$

In order to show the existence and uniqueness of the solution for (3.1), we show the equivalence of M_2 and \mathcal{L}_2 , where the equivalence on this norm is dependent on the viscosity ν and the penalty parameter σ . In order to provide such validity of equivalence, we need to define the proper solution spaces

$$\mathcal{W}_1 = L^2(\Omega)^{n^2} \times H_0^1(\Omega)^n \times L_0^2(\Omega) \quad \text{and} \quad \mathcal{W} = \mathcal{V}_0 \times H_0^1(\Omega)^n \times (H^1(\Omega) \cap L_0^2(\Omega)),$$

where $\mathcal{V}_0 = \{\mathbf{V} \in H^1(\Omega)^{n^2} : \mathbf{n} \times \mathbf{V} = \mathbf{0} \text{ on } \partial\Omega\}$. From now on, we assume that the penalty parameter σ is in the range $0 < \sigma \leq 1$ because the small values of σ are required for good control (see Section 5). The following theorem shows the equivalence of M_1 and \mathcal{L}_1 .

Theorem 3.1. *There are two positive constants C_1 and C_2 dependent on ν and σ such that for any $(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \in \mathcal{W}_1 \times \mathcal{W}_1$, we have*

$$C_1 M_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \leq \mathcal{L}_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0}) \leq C_2 M_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q). \quad (3.4)$$

Proof. One may easily show the upper bound using the triangle, Cauchy-Schwarz inequalities and definition of H^{-1} norm. Also, by modifying the proof of Theorem 3.1 in [12], one may have

$$\begin{aligned} & \|\mathbf{U}\|^2 + \|\nabla \mathbf{u}^t\|^2 + \|\mathbf{V}\|^2 + \nu^2 \|\nabla \mathbf{v}^t\|^2 + \frac{1}{\sigma} \|\mathbf{v}\|^2 \\ & \leq C \max\left\{1 + \frac{1}{\nu^2}, \sigma^2 + \sigma \nu^2\right\} \mathcal{L}_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0}). \end{aligned} \quad (3.5)$$

Combining Lemma 3.1 and (3.5) yields the low bound of (3.4) with a positive constant C_1 dependent on ν and σ . \square

For the norm equivalence of $\mathcal{L}_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0})$ with $M_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q)$, we need to establish the H^2 -regularity estimates of the following coupled equations:

$$\left\{ \begin{array}{ll} -\Delta \mathbf{u} + \nabla p / \nu + \frac{\mathbf{v}}{\sigma \nu} = \mathbf{0} & \text{in } \Omega, \\ \Delta(\mathbf{v} - \mathbf{u}) + \nabla q = \nabla \hat{\mathbf{u}}^t & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \nabla \cdot (\mathbf{v} - \mathbf{u}) = 0 & \text{in } \Omega, \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \mathbf{u} = \mathbf{0} \quad \text{on } \partial\Omega, \\ \mathbf{v} = \mathbf{0} \quad \text{on } \partial\Omega. \end{array} \right. \quad (3.6)$$

For uncoupled Stokes-like equations, its H^2 -regularity was provided in [21] by appealing ADN theory (see [2]). Now, in Appendix we provide the H^2 -regularity of (3.6) for a reader's purpose.

Proposition 3.1. *Suppose that the domain Ω is a bounded convex polyhedron or has $C^{1,1}$ boundary. Then, for $\mathbf{u}, \mathbf{v} \in H^2(\Omega)^n \cap H_0^1(\Omega)^n$ and $p, q \in H^1(\Omega)$, the coupled Stokes equations (3.6) satisfies the H^2 regularity estimate*

$$\begin{aligned} & \|\mathbf{u}\|_2 + \|\mathbf{v} - \mathbf{u}\|_2 + \|p\|_1 + \|q\|_1 \\ & \leq C_r \left(\left\| v\Delta\mathbf{u} - \nabla p - \frac{\mathbf{v}}{\sigma} \right\| + \|\Delta(\mathbf{v} - \mathbf{u}) + \nabla q\| + \|\nabla \cdot \mathbf{u}\|_1 + \|\nabla \cdot (\mathbf{v} - \mathbf{u})\|_1 \right), \end{aligned} \tag{3.7}$$

where C_r depends on v, σ and Ω .

Proof. See Appendix 6. □

We summarize some results from [17] for the next theorem. The following inequalities are from Theorems 3.7-3.9 and Lemmas 3.2, 3.6 in [17], respectively. Assume that the domain Ω is a bounded convex polyhedron or has $C^{1,1}$ boundary. Then for any vector \mathbf{v} in either $H_0(\text{div}) \cap H(\mathbf{curl})$ or $H(\text{div}) \cap H_0(\mathbf{curl})$ it follows that

$$\|\mathbf{v}\|_1^2 \leq C \left(\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2 + \|\nabla \times \mathbf{v}\|^2 \right). \tag{3.8}$$

If, in addition, the domain is simply connected, then

$$\|\mathbf{v}\|_1^2 \leq C \left(\|\nabla \cdot \mathbf{v}\|^2 + \|\nabla \times \mathbf{v}\|^2 \right). \tag{3.9}$$

Now it is ready to prove the norm equivalence for the least-squares functional \mathcal{L}_2 with the norm M_2 .

Theorem 3.2. *Assume that the domain Ω is a bounded convex polyhedron or has $C^{1,1}$ boundary. Then there are two constants C_1 and C_2 dependent on σ and v such that for any $(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \in \mathcal{W} \times \mathcal{W}$, we have*

$$C_1 M_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \leq \mathcal{L}_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0}) \leq C_2 M_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q). \tag{3.10}$$

Proof. The upper bound in (3.10) is straightforward from the triangle and Cauchy-Schwarz inequalities. Since $\mathcal{W} \subset \mathcal{W}_1$, we have $\mathcal{L}_1 \leq \mathcal{L}_2$ on $\mathcal{W} \times \mathcal{W}$. By Theorem 3.1,

$$M_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q) \leq C \mathcal{L}_1(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0}) \leq C \mathcal{L}_2(\mathbf{U}, \mathbf{u}, p, \mathbf{V}, \mathbf{v}, q; \mathbf{0}).$$

From the inequality (3.8) and the standard Poincaré-Friedrichs inequality, we have

$$\begin{aligned} & \|\mathbf{U}\|_1^2 + \|p\|_1^2 + \|\mathbf{V}\|_1^2 + \|q\|_1^2 \\ & \leq C \left(\|\mathbf{U}\|^2 + \|(\nabla \cdot \mathbf{U})^\dagger\|^2 + \|\nabla \times \mathbf{U}\|^2 + \|\nabla p\|^2 \right. \\ & \quad \left. + \|\mathbf{V}\|^2 + \|(\nabla \cdot \mathbf{V})^\dagger\|^2 + \|\nabla \times \mathbf{V}\|^2 + \|\nabla q\|^2 \right). \end{aligned}$$

It thus suffices to show that

$$\begin{aligned} & C(\|(\nabla \cdot \mathbf{U})^t\|^2 + \|\nabla p\|^2 + \|(\nabla \cdot \mathbf{V})^t\|^2 + \|\nabla q\|^2) \\ & \leq \left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \right\|^2 + |\operatorname{tr} \mathbf{U}|_1^2 + \|\nabla \times \mathbf{U}\|^2 + \|\mathbf{U} + \mathbf{V} - \nu \nabla \mathbf{v}^t\|^2 \\ & \quad + \|(\nabla \cdot \mathbf{V})^t + \nabla q\|^2 + |\operatorname{tr} \mathbf{V}|_1^2 + \|\nabla \times \mathbf{V}\|^2. \end{aligned} \quad (3.11)$$

If (3.11) is satisfied for simply connected Ω , then, by the similar arguments in the proof of Theorem 3.7 in [17], it is also satisfied for Ω whose boundary $\partial\Omega$ is $C^{1,1}$. Hence it is enough to assume that the domain Ω is simply connected with connected boundary. Also we will prove (3.11) only for three dimensional case because its proof can be reduced to two dimensional case. Since \mathbf{U} and \mathbf{V} are in \mathcal{V}_0 , there exist \mathbf{r} , \mathbf{w} , Φ and Ψ satisfying Corollary 3.4 in [17] such that

$$\mathbf{U} = \nabla \mathbf{r}^t + \nabla \times \Phi, \quad \text{and} \quad \mathbf{U} + \mathbf{V} = \nabla \mathbf{w}^t + \nabla \times \Psi, \quad (3.12)$$

with

$$\Delta \mathbf{r} = (\nabla \cdot \mathbf{U})^t \quad \text{and} \quad \Delta(\mathbf{w} - \mathbf{r}) = (\nabla \cdot \mathbf{V})^t.$$

By taking the curl of both sides of decomposition (3.12), it is easy to see that

$$\|\Delta \Phi\| = \|\nabla \times \mathbf{U}\| \quad \text{and} \quad \|\Delta(\Psi - \Phi)\| = \|\nabla \times \mathbf{V}\|. \quad (3.13)$$

Hence,

$$\begin{aligned} & \|(\nabla \cdot \mathbf{U})^t\|^2 + \|\nabla p\|^2 + \|(\nabla \cdot \mathbf{V})^t\|^2 + \|\nabla q\|^2 \\ & = \|\Delta \mathbf{r}\|^2 + \|\nabla p\|^2 + \|\Delta(\mathbf{w} - \mathbf{r})\|^2 + \|\nabla q\|^2 \\ & \leq C_r \left(\left\| \nu \Delta \mathbf{r} - \nabla p - \frac{\mathbf{w}}{\sigma} \right\|^2 + \|\nabla \cdot \mathbf{r}\|_1^2 + \|\Delta(\mathbf{w} - \mathbf{r}) + \nabla q\|^2 + \|\nabla \cdot (\mathbf{w} - \mathbf{r})\|_1^2 \right) \end{aligned} \quad (3.14)$$

$$\leq C \left(\left\| \nu \Delta \mathbf{r} - \nabla p - \frac{\mathbf{w}}{\sigma} \right\|^2 + \|\nabla \cdot \mathbf{r}\|_1^2 + \|\Delta(\mathbf{w} - \mathbf{r}) + \nabla q\|^2 + \|\nabla \cdot (\mathbf{w} - \mathbf{r})\|_1^2 \right) \quad (3.15)$$

$$\begin{aligned} & \leq C \left(\left\| \nu \Delta \mathbf{r} - \nabla p - \frac{\mathbf{w}}{\sigma} \right\|^2 + \|\nabla \cdot \mathbf{r} + \operatorname{tr} \nabla \times \Phi\|_1^2 + \|\Delta \Phi\|^2 \right. \\ & \quad \left. + \|\nu \Delta(\mathbf{w} - \mathbf{r}) + \nabla q\|^2 + \|\nabla \cdot (\mathbf{w} - \mathbf{r}) + \operatorname{tr} \nabla \times (\Psi - \Phi)\|_1^2 + \|\Delta(\Psi - \Phi)\|^2 \right) \end{aligned} \quad (3.16)$$

$$\begin{aligned} & = C \left(\left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{w}}{\sigma} \right\|^2 + |\operatorname{tr} \mathbf{U}|_1^2 + \|\nabla \times \mathbf{U}\|^2 \right. \\ & \quad \left. + \|(\nabla \cdot \mathbf{V})^t + \nabla q\|^2 + |\operatorname{tr} \mathbf{V}|_1^2 + \|\nabla \times \mathbf{V}\|^2 \right). \end{aligned} \quad (3.17)$$

The inequalities (3.14), (3.15) and (3.16) are from (3.7), Lemma 3.1 and Lemma 3.2 in [12] respectively. The equality (3.17) is from (3.13). Using triangle inequality, we have

$$\left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{w}}{\sigma} \right\|^2 \leq C \left(\left\| \nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \right\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 \right).$$

Using Poincaré-Friedrichs inequality and (3.12), it follows that

$$\|\mathbf{v} - \mathbf{w}\| \leq C \|\nabla \mathbf{v}^t - \nabla \mathbf{w}^t\|^2 = C \left(\|\nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V} + \nabla \times \boldsymbol{\Psi}\|^2 \right). \quad (3.18)$$

Then applying triangle inequality and (3.9) to the right hand side of (3.18) with $\nabla \times \boldsymbol{\Psi}$, and using (3.13), we have

$$\begin{aligned} \|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V} + \nabla \times \boldsymbol{\Psi}\|^2 &\leq C (\|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V}\|^2 + \|\nabla \times \boldsymbol{\Psi}\|_1^2) \\ &\leq C (\|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V}\|^2 + \|\nabla \cdot \nabla \times \boldsymbol{\Psi}\|^2 + \|\nabla \times \nabla \times \boldsymbol{\Psi}\|^2) \\ &= C (\|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V}\|^2 + \|\Delta \boldsymbol{\Psi}\|) \\ &= C (\|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V}\|^2 + \|\nabla \times (\mathbf{U} + \mathbf{V})\|^2) \\ &\leq C (\|\nu \nabla \mathbf{v}^t - \mathbf{U} - \mathbf{V}\|^2 + \|\nabla \times \mathbf{U}\|^2 + \|\nabla \times \mathbf{V}\|^2). \end{aligned}$$

This proves (3.11) for simply connected Ω . Hence, we have the conclusion. \square

4 Finite element approximations

For the finite element approximation, let \mathcal{T}_h be a partition of the Ω into finite elements with $h = \max\{\text{diam}(K) : K \in \mathcal{T}_h\}$. Assume that the triangulation \mathcal{T}_h is a quasi-uniform, i.e., it is regular and satisfies the inverse assumption (see [11]). Let $\mathcal{W}^h := \mathcal{V}^h \times \mathcal{U}^h \times \mathcal{Q}^h$ be a finite dimensional subspace of \mathcal{W} with the following approximation properties : for any

$$(\mathbf{U}, \mathbf{u}, p) \in \mathcal{W} \cap (H^{r+1}(\Omega)^{n^2} \times H^{s+1}(\Omega)^n \times H^{k+1}(\Omega)) \quad (r, s, k \geq 1),$$

there exist a constant C and a pair $(\mathbf{U}_h, \mathbf{u}_h, p_h) \in \mathcal{W}^h$ such that

$$\inf_{\mathbf{U}_h \in \mathcal{V}^h} \{ \|\mathbf{U} - \mathbf{U}_h\|_0 + h \|\mathbf{U} - \mathbf{U}_h\|_1 \} \leq Ch^{r+1} \|\mathbf{U}\|_{r+1}, \quad (4.1)$$

$$\inf_{\mathbf{u}_h \in \mathcal{U}^h} \{ \|\mathbf{u} - \mathbf{u}_h\|_0 + h \|\mathbf{u} - \mathbf{u}_h\|_1 \} \leq Ch^{s+1} \|\mathbf{u}\|_{s+1}, \quad (4.2)$$

$$\inf_{p_h \in \mathcal{Q}^h} \{ \|p - p_h\|_0 + h \|p - p_h\|_1 \} \leq Ch^{k+1} \|p\|_{k+1}. \quad (4.3)$$

For convenience, let the state variables and adjoint variables be $\Phi := \{\mathbf{U}, \mathbf{u}, p\} \in \mathcal{W}$ and $\Lambda := \{\mathbf{V}, \mathbf{v}, q\} \in \mathcal{W}$, respectively. Then, the variational problem equivalent to (3.1) is to find $(\Phi, \Lambda) \in \mathcal{W} \times \mathcal{W}$ such that

$$\mathcal{A}((\Phi, \Lambda), (\tilde{\Phi}, \tilde{\Lambda})) = F((\tilde{\Phi}, \tilde{\Lambda}); \hat{\mathbf{u}}) \quad \forall (\tilde{\Phi}, \tilde{\Lambda}) \in \mathcal{W} \times \mathcal{W}, \quad (4.4)$$

where

$$\begin{aligned} & \mathcal{A}((\Phi, \Lambda), (\tilde{\Phi}, \tilde{\Lambda})) \\ &= \left(\nu(\nabla \cdot \mathbf{U})^t - \nabla p - \frac{\mathbf{v}}{\sigma} \nu(\nabla \cdot \tilde{\mathbf{U}})^t - \nabla \tilde{p} - \frac{\tilde{\mathbf{v}}}{\sigma} \right) \\ & \quad + \nu^2(\mathbf{U} - \nabla \mathbf{u}^t, \tilde{\mathbf{U}} - \nabla \tilde{\mathbf{u}}^t) + \nu^2(\nabla \cdot \mathbf{u}, \nabla \cdot \tilde{\mathbf{u}}) + \nu^2(\nabla \times \mathbf{U}, \nabla \times \tilde{\mathbf{U}}) \\ & \quad + \nu^2(\nabla \operatorname{tr}(\mathbf{U}), \nabla \operatorname{tr}(\tilde{\mathbf{U}})) + \frac{\nu^2}{\sigma^2}((\nabla \cdot \mathbf{V})^t + \nabla q, (\nabla \cdot \tilde{\mathbf{V}})^t + \nabla \tilde{q}) \\ & \quad + \frac{1}{\sigma^2}(\mathbf{V} - \nu \nabla \mathbf{v}^t + \mathbf{U}, \tilde{\mathbf{V}} - \nu \nabla \tilde{\mathbf{v}}^t + \tilde{\mathbf{U}}) + \frac{\nu^2}{\sigma^2}(\nabla \cdot \mathbf{v}, \nabla \cdot \tilde{\mathbf{v}}) + \frac{\nu^2}{\sigma^2}(\nabla \times \mathbf{V}, \nabla \times \tilde{\mathbf{V}}) \\ & \quad + \frac{\nu^2}{\sigma^2}(\nabla \operatorname{tr}(\mathbf{V}), \nabla \operatorname{tr}(\tilde{\mathbf{V}})), \end{aligned}$$

and

$$\mathcal{F}((\tilde{\Phi}, \tilde{\Lambda}); \hat{\mathbf{u}}) = \frac{1}{\sigma^2}(\nabla \hat{\mathbf{u}}, \tilde{\mathbf{V}} - \nu \nabla \tilde{\mathbf{v}}^t + \tilde{\mathbf{U}}).$$

The corresponding finite element discretization of (4.4) is to find $(\Phi_h, \Lambda_h) \in \mathcal{W}^h \times \mathcal{W}^h$ satisfying

$$\mathcal{A}((\Phi_h, \Lambda_h), (\tilde{\Phi}_h, \tilde{\Lambda}_h)) = \mathcal{F}((\tilde{\Phi}_h, \tilde{\Lambda}_h); \hat{\mathbf{u}}) \quad \forall (\tilde{\Phi}_h, \tilde{\Lambda}_h) \in \mathcal{W}^h \times \mathcal{W}^h. \tag{4.5}$$

Then, the established ellipticity and continuity in a product H^1 norm may yield the following optimal discretization error estimates in the finite element space $\mathcal{W}^h \times \mathcal{W}^h$.

Proposition 4.1. *Let (Φ, Λ) be the solution of the minimization of \mathcal{L}_2 over $\mathcal{W} \times \mathcal{W}$ and (Φ_h, Λ_h) the unique minimizer of \mathcal{L}_2 over $\mathcal{W}_h \times \mathcal{W}_h$. Then*

$$M_2(\Phi - \Phi_h, \Lambda - \Lambda_h) \leq C \inf_{(\tilde{\Phi}_h, \tilde{\Lambda}_h) \in \mathcal{W}_h \times \mathcal{W}_h} M_2(\Phi - \tilde{\Phi}_h, \Lambda - \tilde{\Lambda}_h). \tag{4.6}$$

Proof. Theorem 3.2, the orthogonality of the error $(\Phi - \tilde{\Phi}_h, \Lambda - \tilde{\Lambda}_h)$ to $\mathcal{W}_h \times \mathcal{W}_h$ with respect to the bilinear form $\mathcal{A}(\cdot, \cdot)$, and the Cauchy-Schwarz inequality imply (4.6). \square

Theorem 4.1. *Assume that $(\Phi, \Lambda) \in [\mathcal{W} \cap (H^{s+1}(\Omega)^{n^2} \times H^{s+1}(\Omega)^n \times H^{s+1}(\Omega))]^2$, where $s \geq 1$, is the solution of the minimization problem for \mathcal{L}_2 and (Φ_h, Λ_h) is the unique minimizer of \mathcal{L}_2 over $\mathcal{W}_h \times \mathcal{W}_h$. Then*

$$\begin{aligned} & M_2(\Phi - \Phi_h, \Lambda - \Lambda_h) \\ & \leq Ch^{2s} \left(\|\mathbf{U}\|_{s+1}^2 + \|\mathbf{u}\|_{s+1}^2 + \|p\|_{s+1}^2 + \|\mathbf{V}\|_{s+1}^2 + \|\mathbf{v}\|_{s+1}^2 + \|q\|_{s+1}^2 \right). \end{aligned}$$

Proof. It can be induced by Proposition 4.1 and the approximation properties (4.1)-(4.3). \square

5 Numerical experimentations

For numerical implementations corresponding to bilinear form (4.5), we take the unit square $\Omega := (0,1) \times (0,1) \subset \mathbb{R}^2$ as a computational domain.

We use the single approximating space of continuous piecewise linear functions on a triangulations \mathcal{T}_h with mesh size h for the approximations of all unknowns to represent the matrix corresponding to (4.5). Let $\mathcal{W}^h := \mathcal{V}^h \times \mathcal{U}^h \times \mathcal{Q}^h$ be such a finite element space. Let us choose bases $\{\mathcal{V}_j(\mathbf{x})\}_{j=1}^{4J}$, $\{\mathcal{U}_j(\mathbf{x})\}_{j=1}^{2J}$, and $\{\mathcal{Q}_j(\mathbf{x})\}_{j=1}^J$ for \mathcal{V}^h , \mathcal{U}^h and \mathcal{Q}^h , respectively, so that one may have

$$\begin{aligned} \mathbf{U}^h &= \sum_{j=1}^{4J} U_j \mathcal{V}_j, & \mathbf{u}^h &= \sum_{j=1}^{2J} u_j \mathcal{U}_j, & p^h &= \sum_{j=1}^J p_j \mathcal{Q}_j, \\ \mathbf{V}^h &= \sum_{j=1}^{4J} V_j \mathcal{V}_j, & \mathbf{v}^h &= \sum_{j=1}^{2J} v_j \mathcal{U}_j, & q^h &= \sum_{j=1}^J q_j \mathcal{Q}_j \end{aligned}$$

for some sets of coefficients $\{U\}_{j=1}^{4J}$, $\{u\}_{j=1}^{2J}$, $\{p\}_{j=1}^J$, $\{V\}_{j=1}^{4J}$, $\{v\}_{j=1}^{2J}$, and $\{q\}_{j=1}^J$ that are determined by solving (5.1). Hence the discrete problem has a matrix equation,

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{B}^t \\ \mathbf{B} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \vec{\Phi} \\ \vec{\Lambda} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{F}}_1 \\ \vec{\mathbf{F}}_2 \end{bmatrix}, \tag{5.1}$$

where

$$\mathbf{A}_1 = \begin{bmatrix} \mathbb{K}_1 & \mathbf{C}_1^t & \mathbf{C}_2^t \\ \mathbf{C}_1 & \mathbb{K}_2 & \mathbf{0} \\ \mathbf{C}_2 & \mathbf{0} & \mathbb{K}_3 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} \mathbb{K}_4 & \frac{1}{\sigma^2 v} \mathbf{C}_1^t & -\frac{v}{\sigma^2} \mathbf{C}_2^t \\ \frac{1}{\sigma^2 v} \mathbf{C}_1 & \mathbb{K}_5 & \mathbf{0} \\ -\frac{v}{\sigma^2} \mathbf{C}_2 & \mathbf{0} & \frac{v^2}{\sigma^2} \mathbb{K}_3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{C}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{C}_4 & \mathbf{0} & \mathbf{C}_5 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

and

$$\vec{\Phi} = \begin{bmatrix} \vec{U} \\ \vec{u} \\ \vec{p} \end{bmatrix}, \quad \vec{\Lambda} = \begin{bmatrix} \vec{V} \\ \vec{v} \\ \vec{q} \end{bmatrix}, \quad \vec{\mathbf{F}}_1 = \begin{bmatrix} \vec{g}_1 \\ \vec{0} \\ \vec{0} \end{bmatrix}, \quad \vec{\mathbf{F}}_2 = \begin{bmatrix} \vec{g}_1 \\ \vec{g}_2 \\ \vec{0} \end{bmatrix}.$$

In (5.1), we have $\vec{U} = (U_1, \dots, U_{4J}), \dots, \vec{q} = (q_1, \dots, q_J)$,

$$\begin{aligned} (\mathbb{K}_1)_{ij} &= v^2 ((\nabla \cdot \mathcal{V}_j)^t, (\nabla \cdot \mathcal{V}_i)^t) + (v^2 + \frac{1}{\sigma^2}) (\mathcal{V}_j, \mathcal{V}_i) \\ &\quad + v^2 (\nabla \times \mathcal{V}_j, \nabla \times \mathcal{V}_i) + v^2 (\nabla \text{tr}(\mathcal{V}_j), \nabla \text{tr}(\mathcal{V}_i)) && \text{for } i, j = 1, \dots, 4J, \\ (\mathbb{K}_2)_{ij} &= v^2 (\nabla \mathcal{U}_j^t, \nabla \mathcal{U}_i^t) + v^2 (\nabla \cdot \mathcal{U}_j, \nabla \cdot \mathcal{U}_i) && \text{for } i, j = 1, \dots, 2J, \\ (\mathbb{K}_3)_{ij} &= (\nabla \mathcal{Q}_j, \nabla \mathcal{Q}_i) && \text{for } i, j = 1, \dots, J, \end{aligned}$$

$$\begin{aligned}
(\mathbb{K}_4)_{ij} &= \frac{\nu^2}{\sigma^2} ((\nabla \cdot \mathcal{V}_j)^t, (\nabla \cdot \mathcal{V}_i)^t) + \frac{1}{\sigma^2} (\mathcal{V}_j, \mathcal{V}_i) \\
&\quad + \frac{\nu^2}{\sigma^2} (\nabla \times \mathcal{V}_j, \nabla \times \mathcal{V}_i) + \frac{\nu^2}{\sigma^2} (\nabla \operatorname{tr}(\mathcal{V}_j), \nabla \operatorname{tr}(\mathcal{V}_i)) \quad \text{for } i, j = 1, \dots, 4J, \\
(\mathbb{K}_5)_{ij} &= \frac{1}{\sigma^2} (\mathcal{U}_j, \mathcal{U}_i) + \frac{\nu^2}{\sigma^2} (\nabla \mathcal{U}_j^t, \nabla \mathcal{U}_i^t) + \frac{\nu^2}{\sigma^2} (\nabla \cdot \mathcal{U}_j, \nabla \cdot \mathcal{U}_i) \quad \text{for } i, j = 1, \dots, 2J, \\
(\mathbf{C}_1)_{ij} &= -\nu (\mathcal{V}_j, \nabla \mathcal{U}_i^t) \quad \text{for } i = 1, \dots, 2J, j = 1, \dots, 4J, \\
(\mathbf{C}_2)_{ij} &= -\nu ((\nabla \cdot \mathcal{V}_j)^t, \nabla \mathcal{Q}_i) \quad \text{for } i = 1, \dots, J, j = 1, \dots, 4J, \\
(\mathbf{C}_3)_{ij} &= \frac{1}{\sigma^2} (\mathcal{V}_j, \mathcal{V}_i) \quad \text{for } i, j = 1, \dots, 4J, \\
(\mathbf{C}_4)_{ij} &= -\frac{\nu}{\sigma} ((\nabla \cdot \mathcal{V}_j)^t, \mathcal{U}_i) - \frac{\nu}{\sigma^2} (\mathcal{V}_j, \nabla \mathcal{U}_i^t) \quad \text{for } i = 1, \dots, 2J, j = 1, \dots, 4J, \\
(\mathbf{C}_5)_{ij} &= \frac{1}{\sigma} (\nabla \mathcal{Q}_j, \mathcal{U}_i) \quad \text{for } i = 1, \dots, 2J, j = 1, \dots, J,
\end{aligned}$$

and

$$\vec{g}_1 = \frac{1}{\sigma^2} (\nabla \hat{\mathbf{u}}^t, \mathcal{V}_j) \quad \text{for } j = 1, \dots, 4J, \quad \vec{g}_2 = -\frac{\nu}{\sigma^2} (\nabla \hat{\mathbf{u}}^t, \nabla \mathcal{U}_j) \quad \text{for } j = 1, \dots, 2J.$$

We note that the matrix in (5.1) from \mathcal{L}_2 functional is symmetric and positive definite. The positivity is dependent on the viscosity ν and the penalty parameter σ . The implementation of a model problem shows the given target velocity $\hat{\mathbf{u}}$ and the velocity flux $\nabla \hat{\mathbf{u}}^t$ can be reached by the finite element solution \mathbf{u}^h and \mathbf{U}^h , respectively, as $h \rightarrow 0$. Also, to show the role of the penalty parameter σ , we report the L^2 errors between the target states $\hat{\mathbf{u}}$, $\nabla \hat{\mathbf{u}}^t$ and the controlled states \mathbf{u}^h , \mathbf{U}^h as $\sigma \rightarrow 0$.

Since the product H^1 -norm equivalence of \mathcal{L}_2 -functional is shown in Theorem 3.2, it is possible to use the multigrid V -cycle preconditioner to solve the linear system (5.1). For terminating the $V(1,1)$ -cycle with the Gauss-Seidel smoothing iteration, the relative residual tolerance $\|R_m\| / \|R_0\| < \epsilon := 10^{-5}$ is used with maximum iteration 300, where R_m is the m^{th} residual of the system (5.1). For the linear system (5.1), the five-points Gaussian quadrature is used on each triangle for all integrals of \vec{g}_ℓ ($\ell = 1, 2$) and set to be zeros as the initial state variables and the adjoint variables for our computations. As a target state, we take the example in [16] as $\hat{\mathbf{u}}(x, y) = (\hat{u}_1(x, y), \hat{u}_2(x, y))$, where

$$\hat{u}_1(x, y) = \frac{1}{27} (\phi(x)\phi'(y)) \quad \text{and} \quad \hat{u}_2(x, y) = -\frac{1}{27} (\phi'(x)\phi(y)) \quad (5.2)$$

with $\phi(z) = (1 - \cos(3\pi z))(1 - z)^2$. With this divergence free $\hat{\mathbf{u}}$, we examine L^2 errors between the finite element solutions \mathbf{U}^h , \mathbf{u}^h and the target variables $\nabla \hat{\mathbf{u}}^t$, $\hat{\mathbf{u}}$ respectively, varying the penalty parameter σ for fixed $\nu = 1$. The penalty parameter is taken as $\sigma \leq 1$ because small value of σ yields a good control result (see [13, 16, 27]).

To show the effects of σ as $\sigma \rightarrow 0$, we report the L^2 errors between the target variables and the controlled variables, and the L^2 norm of control \mathbf{f}^h with the value of the functional

Table 1: Numerical results.

h	σ	$\ \mathbf{U}^h - \nabla \hat{\mathbf{u}}^t\ $	$\ \mathbf{u}^h - \hat{\mathbf{u}}\ $	$\ \mathbf{f}^h\ $	$\mathcal{J}(\mathbf{U}^h, \mathbf{f}^h)$
$\frac{1}{32}$	1	$8.0662e-1$	$6.6910e-2$	$6.5519e-2$	$4.3849e-3$
	10^{-1}	$7.5195e-1$	$6.2137e-2$	$6.0841e-1$	$2.0438e-2$
	10^{-2}	$3.3257e-1$	$2.7469e-2$	$2.6409e+0$	$3.5250e-2$
	10^{-3}	$3.8343e-2$	$4.4055e-3$	$3.0704e+0$	$4.7232e-3$
	10^{-4}	$8.9246e-3$	$2.2023e-3$	$5.1072e+0$	$1.3066e-3$
	10^{-5}	$7.3180e-3$	$2.0383e-3$	$2.8993e+1$	$4.2299e-3$
$\frac{1}{64}$	1	$8.0654e-1$	$6.6898e-2$	$6.6543e-2$	$3.2747e-1$
	10^{-1}	$7.5807e-1$	$6.2503e-2$	$6.2173e-1$	$3.0666e-1$
	10^{-2}	$4.3348e-1$	$3.4767e-2$	$3.4510e+0$	$1.5350e-1$
	10^{-3}	$6.9981e-2$	$5.7931e-3$	$5.4945e+0$	$1.7543e-2$
	10^{-4}	$7.8282e-3$	$9.7505e-4$	$6.0307e+0$	$1.8491e-3$
	10^{-5}	$1.9550e-3$	$5.3648e-4$	$9.5394e+0$	$4.5691e-4$

$\mathcal{J}(\mathbf{U}^h, \mathbf{f}^h)$ for the target velocity (5.2). Table 1 shows the L^2 errors between the target variables (flux and velocity) and approximated controlled variables, respectively plus with the L^2 norm of control \mathbf{f} . In Table 1, the values of corresponding objective functional are presented with mesh size $h = 1/32$ or $h = 1/64$. We can see the distances $\|\mathbf{U}^h - \nabla \hat{\mathbf{u}}^t\|$, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$ and the values of objective functional are all decreasing and the magnitude of control \mathbf{f}^h is increasing as $\sigma \rightarrow 0$. Also, in Table 1, it is observed that the L^2 distances of fluxes are always larger than the ones of velocities as predicted by (1.3). Fig. 1 shows the

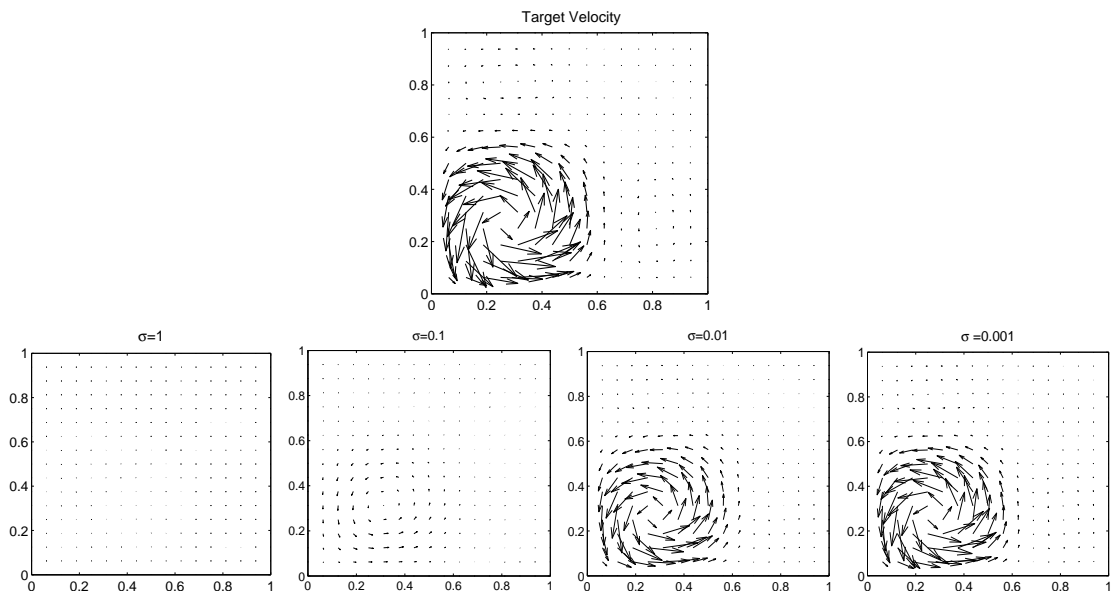


Figure 1: Target flow and controlled flows for $\sigma = 1, 10^{-1}, 10^{-2}, 10^{-3}$ when $h = 1/16$.

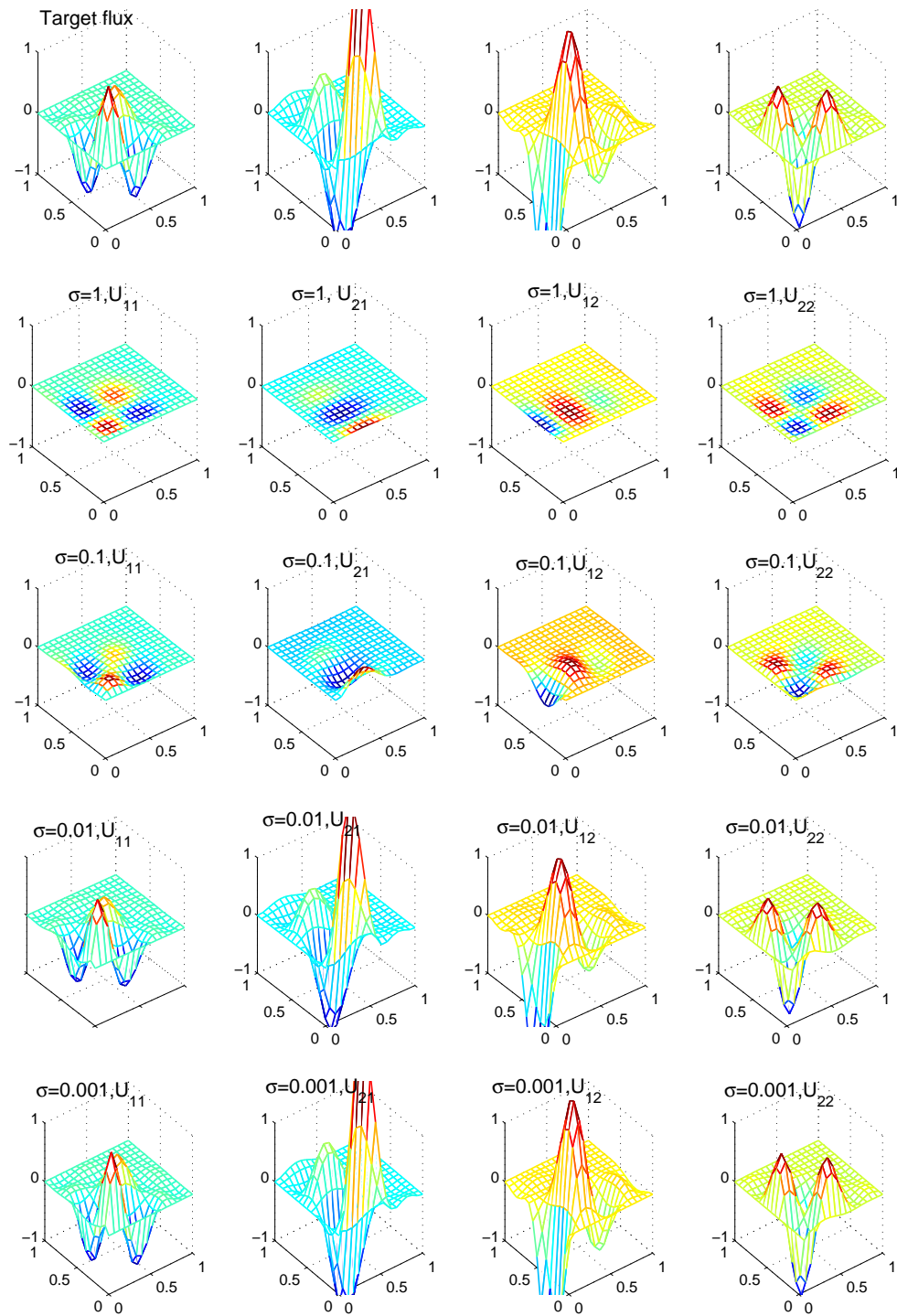


Figure 2: Target velocity flux and controlled fluxes for $\sigma=1, 10^{-1}, 10^{-2}, 10^{-3}$ when $h=1/16$ and $\nu=1$.

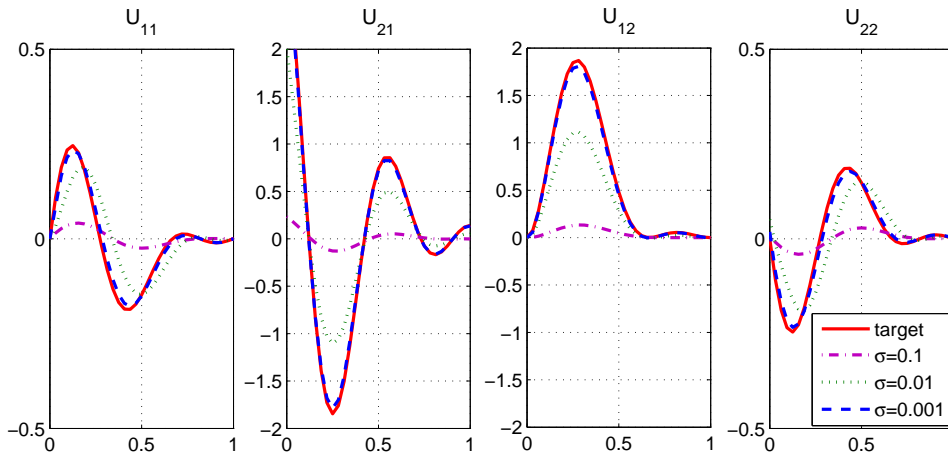


Figure 3: Sectional graphs for the target flux and controlled fluxes at different σ values for fixed x -axis ($x=0.25$) when $h=1/32$ and $\nu=1$.

behaviors of target velocity and controlled velocities when $\sigma = 1, 0.1, 0.01, 0.001$. In these table and figure, we see that the controlled velocity and the controlled flux converge to the target velocity and the target flux, respectively, as $\sigma \rightarrow 0$.

In Fig. 2, it is presented that the elements of the target flux and the elements of the controlled flux U_{11}, U_{21}, U_{12} and U_{22} when $\sigma = 1, 0.1, 0.01, 0.001$. To compare the flux solutions in Fig. 2, sectional each flux element is plotted in the same figure for different σ and fixed x -axis ($x=0.25$) in Fig. 3. In these figures, we can see the velocity flux is also controlled well for small σ values. Table 2 exhibits the values $\|\mathbf{U}^h - \nabla \hat{\mathbf{u}}^t\|$ and $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$ when $\sigma = 10^{-5}$ as mesh size h is decreasing. Through the values of Table 2, we may say that \mathbf{U}^h and \mathbf{u}^h are controlled well for small σ and h . Note that in Table 2, the convergence rates are like $\mathcal{O}(h^2)$ for $\sigma = 10^{-5}$.

Table 2: The L^2 errors $\|\mathbf{U}^h - \nabla \hat{\mathbf{u}}^t\|$ and $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$ when $\sigma = 10^{-5}$.

h	$\ \mathbf{U}^h - \nabla \hat{\mathbf{u}}^t\ $	ρ	$\ \mathbf{u}^h - \hat{\mathbf{u}}\ $	ρ
$\frac{1}{4}$	$4.2600e-1$		$5.7187e-2$	
$\frac{1}{8}$	$1.4140e-1$	1.6	$2.5142e-2$	1.2
$\frac{1}{16}$	$3.1636e-2$	2.2	$7.7234e-3$	1.7
$\frac{1}{32}$	$7.3180e-3$	2.1	$2.0383e-3$	1.9
$\frac{1}{64}$	$1.9550e-3$	1.9	$5.3648e-4$	1.9

6 Concluding remarks

The FOSLS finite element approach for an optimal control problem governed by the Stokes flows is discussed. The objective functional we consider has H^1 seminorm dis-

tance in the velocity control term. After reformulating the system to a first-order constrained optimal control problem, we have the first-order optimality system using the Lagrange multiplier rule. We impose some weights presented in [27] on least-squares functional and then show the main least-squares functional is equivalent to H^1 product norms. We obtain that the discrete algebraic system induced from L^2 norm based functional is symmetric and positive definite. As a computational result, we observed the velocity flux and the velocity which are controlled as $\sigma \rightarrow 0$.

Appendix: Proof of regularity estimates (3.7)

The ADN theory [2] can be used to prove estimates for the coupled optimality equations (3.6) defined on a domain Ω prescribed. The proof is almost the same as the one in [27] so that we provide a concise proof in this section. First, we define the uniform ellipticity of the systems (3.6) following ADN [2] and introduce the conditions that are necessary for the a priori estimates of [2] to hold.

Proof. We provide $n=3$ dimensional case. Let $l = \{l_{ij}\}$, for $1 \leq i, j \leq 2n+2$ and $B = \{B_{\mu j}\}$, for $1 \leq \mu \leq 2n, 1 \leq j \leq 2n+2$ denote the differential operator and boundary operator corresponding to (3.6). Then we have

$$l = \begin{bmatrix} -\Delta & 0 & 0 & \frac{1}{\sigma\nu} & 0 & 0 & \partial_1 & 0 \\ 0 & -\Delta & 0 & 0 & \frac{1}{\sigma\nu} & 0 & \partial_2 & 0 \\ 0 & 0 & -\Delta & 0 & 0 & \frac{1}{\sigma\nu} & \partial_3 & 0 \\ -\Delta & 0 & 0 & \Delta & 0 & 0 & 0 & \partial_1 \\ 0 & -\Delta & 0 & 0 & \Delta & 0 & 0 & \partial_2 \\ 0 & 0 & -\Delta & 0 & 0 & \Delta & 0 & \partial_3 \\ \partial_1 & \partial_2 & \partial_3 & 0 & 0 & 0 & 0 & 0 \\ -\partial_1 & -\partial_2 & -\partial_3 & \partial_1 & \partial_2 & \partial_3 & 0 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \\ v_3 \\ p/\nu \\ q \end{bmatrix},$$

and

$$F = [0 \ 0 \ 0 \ \nabla \hat{u}_1^t \ \nabla \hat{u}_2^t \ \nabla \hat{u}_3^t \ 0 \ 0]^t, \quad B_{\mu j} = \delta_{\mu j},$$

where $\delta_{\mu j}$ is the Kronecker delta.

Following the developments and notations in [2], we assign a system of integer indices $\{s_i\}, s_i \leq 0$, for the equations and $\{t_j\}, t_j \geq 0$, for the unknown functions. Then, for the system (3.6), we choose Sobolev norms based on the scales

$$s_i = 0 \ (1 \leq i \leq 2n), \quad s_{2n+1} = s_{2n+2} = -1$$

for the equations and

$$t_j = 2 \ (1 \leq j \leq 2n), \quad t_{2n+1} = t_{2n+2} = 1$$

for the variables. Next, the principal part $l'(\partial)$ of the interior operator $l(\partial)$ can be chosen by taking any term (i, j) whose order is $(s_i + t_j)$ as

$$l'(\partial) = \begin{bmatrix} -\Delta & 0 & 0 & 0 & 0 & 0 & \partial_1 & 0 \\ 0 & -\Delta & 0 & 0 & 0 & 0 & \partial_2 & 0 \\ 0 & 0 & -\Delta & 0 & 0 & 0 & \partial_3 & 0 \\ -\Delta & 0 & 0 & \Delta & 0 & 0 & 0 & \partial_1 \\ 0 & -\Delta & 0 & 0 & \Delta & 0 & 0 & \partial_2 \\ 0 & 0 & -\Delta & 0 & 0 & \Delta & 0 & \partial_3 \\ \partial_1 & \partial_2 & \partial_3 & 0 & 0 & 0 & 0 & 0 \\ -\partial_1 & -\partial_2 & -\partial_3 & \partial_1 & \partial_2 & \partial_3 & 0 & 0 \end{bmatrix}.$$

Next, replacing ∂_j by ξ_j ($1 \leq j \leq n$), respectively, the determinant of the principal part can be shown as $L(\xi) := \det(l'(\xi)) = |\xi|^{12}$. In general, we can calculate $L(\xi) := (-1)^{n-1} |\xi|^{4n} \neq 0$ ($n = 2, 3$) for all $\xi \neq 0$, where $|\xi|^2 = \xi_1^2 + \dots + \xi_n^2$. Hence the ellipticity of (3.6) is shown so that the *uniform ellipticity condition*

$$A^{-1} |\xi|^{2m} \leq |L(\xi)| \leq A |\xi|^{2m}$$

holds with $m = 2n$ and $A = 1$. It is easy to see that l' also satisfies the Supplementary Condition.

For *complementing boundary conditions*, note that the matrix for the boundary operator consists of $B_{\mu\eta}$ where $1 \leq \mu \leq 2n$, $1 \leq \eta \leq 2n + 2$ ($n = 2, 3$). If we take $r_\mu = -2$ for $\mu = 1, \dots, 2n$, then $B_{\mu\eta} \leq r_\mu + t_\eta$. The principal boundary operator B' is the same as B . At any point \mathbf{x} of $\partial\Omega$, let \mathbf{n} denote the outward unit normal, $\boldsymbol{\tau} \neq \mathbf{0}$ (real) any tangent unit vector to $\partial\Omega$, $\hat{l}(\xi)$ the stencil of the interior operator l , and $\hat{l}'(\xi)$ the stencil of the principal interior operator l' , where $\xi = \boldsymbol{\tau} + \gamma\mathbf{n}$. Note that the stencil matrix corresponding to the boundary operator B is B because it is the constant matrix. It is easy to show the only root of $L(\boldsymbol{\tau} + \gamma\mathbf{n}) = 0$ with positive imaginary part is i with multiplicity $2n = m$. To show the *complementing boundary conditions* we must show that

$$\sum_{\mu=1}^m C_\mu \sum_{\eta=1}^{2n+1} B'_{\mu\eta}(\mathbf{x}, \boldsymbol{\tau} + \gamma\mathbf{n}) \text{adj}(\hat{l}'(\boldsymbol{\tau} + \gamma\mathbf{n}))_{\eta k} \equiv 0 \pmod{M^+(\gamma)}, \tag{A.1}$$

if and only if the constants C_μ all vanish, where $M^+(\gamma) = (\gamma - i)^m$. We will show the condition (A.1) only for $n = 3$ because its proof is similar to the case $n = 2$. Note the inverse of $\hat{l}'(\xi)$ is

$$\hat{l}'^{-1}(\mathbf{x}, \xi) = \begin{bmatrix} \mathcal{M}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathcal{C}_{3 \times 1} & \mathbf{0}_{3 \times 1} \\ \mathcal{M}_{3 \times 3} & -\mathcal{M}_{3 \times 3} & \mathcal{C}_{3 \times 1} & \mathcal{C}_{3 \times 1} \\ \mathcal{C}_{1 \times 3}^t & \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & \mathcal{C}_{1 \times 3}^t & 0 & -1 \end{bmatrix}_{8 \times 8},$$

where

$$\mathcal{M} = \frac{1}{\rho^2} \begin{bmatrix} -(\xi_2^2 + \xi_3^2) & \xi_1 \xi_2 & \xi_3 \xi_1 \\ \xi_1 \xi_2 & -(\xi_1^2 + \xi_3^2) & \xi_2 \xi_3 \\ \xi_3 \xi_1 & \xi_2 \xi_3 & -(\xi_1^2 + \xi_2^2) \end{bmatrix}, \quad \mathcal{C} = \frac{1}{\rho} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}, \quad \rho = |\xi|^2.$$

From now on, we may assume that $\mathbf{n} = (0,0,1)$ and $\boldsymbol{\tau} = (a,b,0)$, where a, b are arbitrary constants satisfying $a^2 + b^2 = 1$. Note that $\boldsymbol{\xi} = \boldsymbol{\tau} + \gamma \mathbf{n} = (a,b,\gamma)$. Then, since

$$L(\boldsymbol{\xi}) = \rho^6, \quad M^+(\gamma) = (\gamma - i)^6, \quad \rho = (1 + \gamma^2),$$

it follows that

$$B'(\mathbf{x}, \boldsymbol{\tau} + \gamma \mathbf{n}) \text{adj}(\hat{l}'(\boldsymbol{\tau} + \gamma \mathbf{n})) = (\gamma^2 + 1)^4 \begin{bmatrix} \hat{\mathcal{M}}_{3 \times 3} & \mathbf{0}_{3 \times 3} & *_{3 \times 1} & \mathbf{0}_{3 \times 1} \\ \hat{\mathcal{M}}_{3 \times 3} & -\hat{\mathcal{M}}_{3 \times 3} & *_{3 \times 1} & *_{3 \times 1} \end{bmatrix}, \quad (\text{A.2})$$

where

$$\hat{\mathcal{M}} = \begin{bmatrix} -(\gamma^2 + b^2) & ab & a\gamma \\ ab & -(\gamma^2 + a^2) & b\gamma \\ a\gamma & b\gamma & -1 \end{bmatrix}$$

and $*_{3 \times 1}$ stands for a matrix that does not affect row independency. Due to the structure of the matrix in (A.2), it is enough to show that

$$(\gamma^2 + 1)^4 (-(\gamma^2 + b^2)C_{1+j} + abC_{2+j} + a\gamma C_{3+j}) = A_1(\gamma - i)^6, \quad (\text{A.3})$$

where $j = 0,3$ has all zero coefficients $C_1 = C_2 = C_3 = C_4 = C_5 = C_6 = 0$ for the row independency of the matrix in (A.1) where A_1 is a polynomial for γ . By straightforward calculations, this can be easily verified. Applying Theorem 10.5 and following remark in [2], we have (3.7). □

Acknowledgments

This work was supported by Korea Research Foundation under grant KRF-2005-070-C00017.

References

- [1] R. A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
- [2] S. Agmon, A. Douglis and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, II, Comm. Pure Appl. Math., 17 (1964), 35-92.
- [3] D. Bedivan and G. Fix, Least-squares methods for optimal shape design problems, Comput. Math. Appl., 30 (1995), 17-25.

- [4] P. Bochev, Least-squares methods for optimal control, *Nonlinear Anal.*, 30 (1997), 237-256.
- [5] P. Bochev and D. Bedivan, Least-squares methods for Navier-Stokes boundary control problems, *Int. J. Comput. Fluid Dyn.*, 9 (1997), 43-58.
- [6] P. Bochev and M. Gunzburger, Analysis of least-squares finite element methods for the Stokes equations, *Math. Comp.*, 63 (1994), 479-506.
- [7] P. Bochev and M. Gunzburger, Least-squares finite element methods for optimization and control problems for the Stokes equations, *Comput. Math. Appl.*, 48 (2004), 1035-1057.
- [8] P. Bochev and M. Gunzburger, Least-squares finite element methods for optimality systems arising in optimization and control problems, *SIAM J. Numer. Anal.*, 43 (2006), 2517-2543.
- [9] P. 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), 990-1009.
- [10] P. Bochev, T. A. Manteuffel and S. F. McCormick, Analysis of velocity-flux least squares principles for the Navier-Stokes equations. Part II, *SIAM J. Numer. Anal.*, 36 (1999), 1125-1144.
- [11] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [12] Z. Cai, T. A. Manteuffel and S. F. McCormick, First-order system least squares for the Stokes equation, with application to linear elasticity, *SIAM J. Numer. Anal.*, 34 (1997), 1727-1741.
- [13] M. Gunzburger, *Perspectives in Flow Control and Optimization*, Adv. Des. Control 5, SIAM, Philadelphia, 2002.
- [14] R. Glowinski and J. He, *On shape optimization and related issues*, Computational methods for optimal design and control (Arlington, VA, 1997), 151-179.
- [15] M. Gunzburger and H.-C. Lee, A penalty/least-squares method for optimal control problems for first-order elliptic systems, *Appl. Math. Comp.*, 107 (2000), 57-75.
- [16] M. Gunzburger and S. Manservigi, Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control, *SIAM J. Numer. Anal.*, 37 (2000), 1481-1512.
- [17] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, New York, 1986.
- [18] J. He, M. Chevalier, R. Glowinski, R. Metcalfe, A. Nordlander and J. Periaux, Drag reduction by active control for flow past cylinders, *Computational mathematics driven by industrial problems (Martina Franca, 1999)*, Lecture Notes in Math., 1739, (Springer, Berlin, 2000), 287-363.
- [19] J. He, R. Glowinski, R. Metcalfe, A. Nordlander and J. Periaux, Active control and drag optimization for flow past a circular cylinder, *J. Comput. Phys.*, 163 (2000), 83-117.
- [20] S. D. Kim, C.-O. Lee, T. A. Manteuffel, S. F. McCormick and O. Rohrlor, First-order system least squares for the Oseen equations, *Numer. Linear Algebra Appl.*, 13 (2006), 523-542.
- [21] 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.
- [22] S. D. Kim and E. Lee, An analysis for the compressible Stokes equations by first-order system of least-squares finite element method, *Numer. Meth. Part. Diff. Eq.*, 17 (2001), 689-699.
- [23] S. D. Kim and B. C. Shin, H^{-1} least-squares method for the velocity-pressure-stress formulation of Stokes equations, *Appl. Numer. Math.*, 40 (2002), 451-465.
- [24] H.-C. Lee and Y. Choi, A least-squares method for optimal control problems for a second-order elliptic systems in two dimensions, *J. Math. Anal. Appl.*, 242 (2000), 105-128.

- [25] J. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, New York, 1971.
- [26] J. Necas, *Equations aux Derivees Partielles*, Presses de l'Universite de Montreal, Quebec, Canada, 1965.
- [27] S. Ryu, H.-C. Lee and S. D. Kim, First-order system least-squares methods for an optimal control problem by the Stokes flow, *SIAM J. Numer. Anal.*, 47 (2009), 1524-1545.