ESAIM: M2AN 51 (2017) 825–849
DOI: 10.1051/m2an/2016040

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

FINITE ELEMENT APPROXIMATION OF DIRICHLET CONTROL USING BOUNDARY PENALTY METHOD FOR UNSTEADY NAVIER–STOKES EQUATIONS

Sivaguru S. Ravindran[1]

Abstract. This paper is concerned with the analysis of the finite element approximations of Dirichlet control problem using boundary penalty method for unsteady Navier–Stokes equations. Boundary penalty method has been used as a computationally convenient approach alternative to Dirichlet boundary control problems governed by Navier–Stokes equations due to its variational properties. Analysis of the mixed Galerkin finite element method applied to the spatial semi-discretization of the optimality system, from which optimal control can be computed, is presented. An optimal $L^\infty(L^2)$ error estimate of the numerical approximations of the optimality system is derived. Feasibility and applicability of the approach are illustrated by numerically solving a canonical flow control problem.

Mathematics Subject Classification. 65M12, 93C20, 76B75, 49J20, 65M60, 93B40, 76D05.

Received February 19, 2016. Revised May 24, 2016. Accepted May 25, 2016.

1. Introduction

Control of fluid flows modeled by Navier–Stokes equations is an important area of research that has undergone major developments both theoretically and computationally in the recent past. It has many applications including in drag reduction, lift enhancement, mixing augmentation and flow induced noise suppression. The complex nonlinearity and high dimensionality of the Navier–Stokes equations pose many theoretical and computational challenges. There is an extensive body of literature devoted to this subject, see [8,11,23] for surveys in this area.

In this paper, we consider mixed Galerkin finite element approximations of Dirichlet control using boundary penalty method for unsteady Navier–Stokes equations. Dirichlet boundary control, while being practical, is considerably more challenging than other controls in every aspect of control development, from analysis to achieving the control objective. The main difficulty with Dirichlet boundary control is that it is non-variational and thus it is nontrivial to identify suitable function space framework without using appropriate boundary lifting. Several approaches have been proposed in the literature to deal with the theoretical and computational difficulties associated with Dirichlet boundary control, see [5,6,13,15,16]. In most of these works, the function space for the controls is H^s with $s \geq \frac{1}{2}$ which makes numerical realizations by finite elements or finite differences

Keywords and phrases. Boundary penalty method, Dirichlet boundary control, Navier–Stokes equations, optimal error estimates, mixed Galerkin finite element, adjoint equations.

[1] Department of Mathematical Sciences, SST 201M, The University of Alabama in Huntsville, Huntsville, AL 35899, USA. ravinds@uah.edu

Article published by EDP Sciences

more involved than if the control space was L^2. In [13] control space is taken to be H^1 (smooth controls) leading to a boundary Laplace equation for the control. In [5], a separation of variable type Dirichlet control of unsteady Navier−Stokes equation is studied. In [6], Dirichlet control of unsteady Navier−Stokes equations is studied. The choice of function space considered there however involves spaces of fractional powers and thus is not convenient computationally. In [15], boundary penalty approach for Dirichlet boundary control of stationary Navier−Stokes equations is studied from theoretical point of view. Error analysis of finite element approximations of penalized formulations of Dirichlet boundary control problems associated with the stationary Navier−Stokes equations is studied in [16]. Starting with the works of [15,16], interest in penalization technique for treating Dirichlet control has increased. It has since been studied in the context of optimal control with other partial differential equations [2–4,21]. In [21], Dirichlet control of unsteady Navier−Stokes type equations related to Soret convection is studied using boundary penalty approach. In particular, convergence of solutions of penalized control problem to the corresponding solutions of the Dirichlet control problem is proved as the penalty parameter approaches to zero.

In this paper, we are concerned with boundary penalty method to treat the Dirichlet boundary velocity control of unsteady Navier−Stokes equations. Boundary penalty method studied here allows one to work with L^2-control space. Unlike other L^2-control space approaches the present method does not lead to optimality conditions that involve normal derivative of the adjoint variable on the boundary. We present an analysis of the mixed Galerkin finite element spatial discretization of the optimality system, from which optimal control can be computed. An optimal $L^\infty(L^2)$ error estimate of the numerical approximations of the optimality system is derived. Feasibility and applicability of the approach are illustrated by numerical experiments.

The remainder of the paper is organized as follows. In Section 2, we review mathematical background materials related to the unsteady Navier−Stokes equations and give a precise description of both Dirichlet boundary control problem and its penalized counterpart. Moreover, we state results regarding existence of optimal solutions for the boundary penalized optimal control problem. In Section 3, we present the mixed Galerkin finite element approximations of optimality system and prove optimal order $L^\infty(L^2(\Omega))$ error estimates for the approximate solutions of the optimality system. In Section 4, we present results from numerical implementation.

2. Preliminaries and continuum optimal control problem

2.1. Preliminaries

In this section, we introduce some notations and collect several facts from functional analysis which will be useful in subsequent sections. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. When finite element approximations are considered, we will assume that Ω is a convex polyhedral domain; otherwise, we will assume that Ω has Lipschitzian boundary Γ. For $p \geq 1$, let $L^p(\Omega)$ denote the linear space of all real Lebesgue measurable functions ϕ and bounded in the usual norm denoted by $\|\phi\|_{L^p(\Omega)}$. The inner product and norm in $L^2(\Omega)$ are denoted by $(\cdot,\cdot)$ and $\|\cdot\|$, respectively. Let $H^s(\Omega)$ be the usual Hilbertian Sobolev space with s derivatives in $L^2(\Omega)$. We denote with $\|\cdot\|_s$ the norm in $H^s(\Omega)$. The closed subspace of functions in $H^1(\Omega)$ with zero trace on Γ will be denoted by $H^1_0(\Omega)$. The closed subspace of functions in $L^2(\Omega)$ with zero mean on Ω will be denoted by $L^2_0(\Omega)$. The dual space of $H^1_0(\Omega)$ will be denoted by $H^{-1}(\Omega)$ and its norm by $\|\cdot\|_{-1}$. The trace space $H^r(\Gamma)$ consists of functions that are the restriction to the boundary of functions in $H^{r+1/2}(\Omega)$, $r > 0$. We denote the norm and inner product for functions in $H^r(\Gamma)$ by $\|\cdot\|_{r,\Gamma}$ and $(\cdot,\cdot)_{r,\Gamma}$, respectively. In the sequel, we denote by boldface letters $\mathbb{R}^2$-valued function spaces such as $\mathbf{L}^2(\Omega) := [L^2(\Omega)]^2$ and $\mathbf{H}^m(\Omega) := [H^m(\Omega)]^2$. We put

$$\mathbf{V} := \left\{\mathbf{v} \in \mathbf{H}^1(\Omega) \;:\; \nabla\cdot\mathbf{v} = 0 \text{ in } \Omega\right\}, \quad \text{and} \quad \mathbf{H} := \left\{\mathbf{v} \in \mathbf{L}^2(\Omega) \;:\; \nabla\cdot\mathbf{v} = 0 \text{ in } \Omega\right\}$$

and

$$\mathbb{V} := \left\{\mathbf{v} \in \mathbf{H}^1_0(\Omega) \;:\; \nabla\cdot\mathbf{v} = 0 \text{ in } \Omega\right\}.$$

We denote the dual of $\mathbf{V}$ by $\mathbf{V}^*$. If we identify $\mathbf{H}$ with its dual $\mathbf{H}^*$, then we get the following continuous and dense embedding:

$$\mathbf{V} \subset \mathbf{H} = \mathbf{H}^* \subset \mathbf{V}^*.$$

For details, concerning these spaces, see for *e.g.* [10].

For a Banach space X, we denote by $L^p(0,T;X)$ the time-space function space endowed with the norm $\|w\|_{L^p(0,T;X)} := \left(\int_0^T \|w\|_X^p \,\mathrm{d}t\right)^{1/p}$ if $1 \le p < \infty$ and $\operatorname{ess\,sup}_{t\in[0,T]} \|w\|_X$ if $p = \infty$. We will often use the abbreviated notation $L^p(X) := L^p(0,T;X)$ for convenience. We also introduce the space $\mathbf{W}(0,T) := \{\mathbf{u} \in L^2(\mathbf{V}) : \ \mathbf{u}_t \in L^2(\mathbb{V}^*)\}$ endowed with the norms $\|\mathbf{u}\|_W := (\|\mathbf{u}\|^2_{L^2(\mathbf{V})} + \|\mathbf{u}_t\|^2_{L^2(\mathbb{V}^*)})^{\frac{1}{2}}$. We further define

$$\mathbf{W}^{\Sigma} := \{\mathbf{g} | \tau(\mathbf{u}) = \mathbf{g} \ \text{ for } \ \mathbf{u} \in \mathbf{W} \ \},$$

where $\tau : \mathbf{W} \to L^2(0,T;\mathbf{H}^{\frac{1}{2}}(\Gamma))$ is the trace operator onto the lateral boundary $\Sigma := \Gamma \times (0,T)$ of the space-time domain $\Omega \times (0,T)$ given by $\tau(\mathbf{u}(t)) = \mathbf{u}(\cdot,t)|_\Gamma$ for a.e in $[0,T]$. We define the norm on $\mathbf{W}^{\Sigma}$ by

$$\|\mathbf{g}\|_{W^{\Sigma}} := \inf_{\substack{\mathbf{u}\in W \\ \tau(\mathbf{u})=\mathbf{g}}} \|\mathbf{u}\|_W.$$

For $\mathbf{g} \in \mathbf{W}^{\Sigma}$, we define by $\mathbf{u}_g$ the unique element in $\mathbf{W}$ that achieves this infimum, see [21].

We end this section by recalling some inequalities that we will use in this paper.

Poincaré-Friedrichs' inequality: For $\mathbf{u} \in \mathbf{H}^1(\Omega)$,

$$\lambda \|\mathbf{u}\|^2 \le \|\nabla \mathbf{u}\|^2 + \|\mathbf{u}\|^2_{0,\Gamma},$$

where $\lambda > 0$ is a constant, see ([18], Thm. 1.9 and [17], Sect. 5.3). For $\epsilon \in (0, Re)$, we have by Poincaré−Friedrich's inequality,

$$\frac{\widehat{C}}{Re}\|\mathbf{v}\|_1^2 \le \frac{1}{Re}\int_\Omega |\nabla \mathbf{v}|^2 \,\mathrm{d}\Omega + \frac{1}{\epsilon}\int_\Gamma |\mathbf{v}|^2 \,\mathrm{d}s \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega). \tag{2.1}$$

Gagliardo−Nirenberg inequality: For $\mathbf{u} \in \mathbf{H}^1(\Omega) \cap L^q(\Omega)$, let $1 \le q \le r < \infty$. Then, for $s = 1 - (q/r)$,

$$\|\mathbf{u}\|_{L^r(\Omega)} \le C \|\mathbf{u}\|^{1-s}_{L^q(\Omega)} \|\mathbf{u}\|_1^s,$$

see, [7, 19]. Notice that with $q = 2$ and $r = 4$, Gagliardo−Nirenberg inequality implies

$$L^2(0,T;X) \cap L^\infty(0,T;Y) \subset L^4(\Omega \times (0,T)),$$

where $X \subset H^1(\Omega)$ and $Y \subset L^2(\Omega)$.

Young's inequality: For any $a, b \ge 0$ and $\epsilon > 0$, and $q, r > 1$

$$ab \le \frac{\epsilon}{q} a^q + \frac{\epsilon^{-\frac{r}{q}}}{r} b^r, \quad \text{with} \quad \frac{1}{q} + \frac{1}{r} = 1.$$

2.2. Continuum optimal control problem

We consider a boundary control problem for unsteady viscous incompressible flow modeled by the Navier−Stokes equations

$$\begin{cases} \partial_t \mathbf{u} - \dfrac{1}{Re} \Delta \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = \mathbf{f} & \text{in } \ \Omega \times [0,T] \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \ \Omega \times [0,T] \end{cases} \tag{2.2}$$

along with boundary and initial conditions

$$\mathbf{u} = \mathbf{g} \quad \text{on} \ \ \Gamma \times [0, T] \tag{2.3}$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \quad \text{in} \ \ \Omega, \tag{2.4}$$

where $\mathbf{u}$ is the velocity, p the pressure, Re the Reynolds number and $\mathbf{g}$ the control.

Letting $\mathcal{U}_{ad} := \mathbf{W}^{\Sigma}$ be the admissible control set for the Dirichlet control problem, we formulate the Dirichlet control problem as follows:

$$(P) \qquad \begin{aligned} &\text{Minimize } \mathcal{J}(\mathbf{u}, \mathbf{g}) = \int_0^T \left[\Theta(\mathbf{u}) + \frac{\gamma}{2} \int_\Gamma |\mathbf{g}|^2 \, \mathrm{d}\Gamma\right] \mathrm{d}t \ \text{ over all } \ \mathbf{g} \in \mathcal{U}_{ad}, \\ &\text{where } (\mathbf{u}, \mathbf{g}) \in \mathbf{W}(0, T) \times \mathcal{U}_{ad} \text{ satisfies} \end{aligned}$$

$$(\partial_t \mathbf{u}, \mathbf{v}) + \frac{1}{Re}(\nabla \mathbf{u}, \nabla \mathbf{v}) + (\mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}), \ \ \forall \mathbf{v} \in \mathbb{V}$$

$$\mathbf{u}|_\Gamma = \mathbf{g} \quad \text{and} \quad \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}(\mathbf{x}).$$

In the cost functional $\mathcal{J}(\mathbf{u}, \mathbf{g})$, $\mathbf{g}$ is the control field and γ is a positive parameter. Moreover, we assume the function $\Theta : \mathbf{H}^1(\Omega) \to \mathbb{R}^+$ appearing in the cost functional is assumed to satisfy the following:

(i) $\Theta(\mathbf{u})$ is convex and lower-semi continuous and;
(ii) $c_1\|\nabla \mathbf{u}\|^2 - c_2\|\mathbf{u}\|_{0,\Gamma}^2 \le \Theta(\mathbf{u}) \le \hat{c}_1\|\nabla \mathbf{u}\|^2 + \hat{c}_2\|\mathbf{u}\|_{0,\Gamma}^2$ for some constants $c_i, \hat{c}_i \in \mathbb{R}^+$, $i = 1, 2$.

The allowed class of functionals of course includes regulation of viscous dissipation function $\Theta_1 = \frac{\delta}{2}[\|\nabla \mathbf{u} + (\nabla \mathbf{u})^T\|^2]$. The other allowed functionals include regulation of kinetic energy in weighted H^1- norm, *i.e.*, $\Theta_2 = \frac{\delta}{2}[\|\nabla \mathbf{u}\|^2 + \|\mathbf{u}\|^2]$ and regulation of square of the vorticity, *i.e.*, $\Theta_3 = \frac{\delta}{2}\|\nabla \times \mathbf{u}\|^2$. The former expression for Θ already satisfies the above assumption. The later can also be shown to satisfy the above assumption (see [15], Lem. 3.3).

In penalized optimal control problem, the Dirichlet boundary control $\mathbf{u}|_\Gamma = \mathbf{g}$ in (2.3) is penalized as

$$-p\mathbf{n} + \frac{1}{Re}\frac{\partial \mathbf{u}}{\partial \mathbf{n}} - \frac{1}{2}(\mathbf{u} \cdot \mathbf{n})\mathbf{u} + \frac{1}{\epsilon}\mathbf{u} = \frac{1}{\epsilon}\mathbf{g} \quad \text{on} \quad \Gamma. \tag{2.5}$$

Moreover, the state and control variables are constrained to satisfy the Navier−Stokes equations in the following weak form

$$\begin{cases} (\partial_t \mathbf{u}, \mathbf{v}) + \frac{1}{Re}(\nabla \mathbf{u}, \nabla \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + \frac{1}{\epsilon}(\mathbf{u}, \mathbf{v})_\Gamma = (\mathbf{f}, \mathbf{v}) + \frac{1}{\epsilon}(\mathbf{g}, \mathbf{v})_\Gamma, \qquad \forall \mathbf{v} \in \mathbf{V} \\ \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}). \end{cases} \tag{2.6}$$

In (2.6), the tri-linear form $c(\cdot, \cdot, \cdot)$ is defined as

$$c(\mathbf{u}, \mathbf{w}, \mathbf{v}) := \frac{1}{2}\left[((\mathbf{u} \cdot \nabla)\mathbf{w}, \mathbf{v}) - ((\mathbf{u} \cdot \nabla)\mathbf{v}, \mathbf{w})\right] = (\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v}) - \frac{1}{2}(\mathbf{v}, \mathbf{w}(\mathbf{u} \cdot \mathbf{n}))_\Gamma$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega)$. Notice that by integration by parts

$$c(\mathbf{u}, \mathbf{w}, \mathbf{v}) = (\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v}) - \frac{1}{2}(\mathbf{v}, \mathbf{w}(\mathbf{u} \cdot \mathbf{n}))_\Gamma$$

and it is clear that $c(\mathbf{u}, \mathbf{v}, \mathbf{v}) = 0$. The existence of a unique solution to the state equation (2.6), see [21], ensures the existence of a control-to-state mapping $\mathbf{g} \mapsto \mathbf{u}(\mathbf{g})$ through (2.6).

The weak formulation of the penalized control problem is defined as follows:

$$(P_\epsilon) \qquad \text{Minimize } \mathcal{J}(\mathbf{u},\mathbf{g}) = \int_0^T \left[\Theta(\mathbf{u}) + \frac{\gamma}{2}\int_\Gamma |\mathbf{g}|^2 \,\mathrm{d}\Gamma\right] \mathrm{d}t \text{ over all } \mathbf{g} \in \mathcal{U}_{ad}^\epsilon,$$
$$\text{where } (\mathbf{u},\mathbf{g}) \in \mathbf{W}(0,T) \times \mathcal{U}_{ad}^\epsilon \text{ satisfies (2.6)},$$

where $\mathcal{U}_{ad}^\epsilon := L^2(0,T;\mathbf{L}^2(\Gamma))$. The existence of the optimal solutions to the penalized control problem $(P)_\epsilon$ is guaranteed by the following theorem in [21].

Theorem 2.1. *Assume* $\epsilon \in (0, Re)$. *Then there exists an optimal solution* $(\mathbf{u},\mathbf{g}) \in \mathbf{W}(0,T)\times\mathcal{U}_{ad}^\epsilon$ *that minimizes* $\mathcal{J}(\mathbf{u},\mathbf{g})$ *subject to* (2.6).

In [21], convergence of the solution of penalized control problem to the solution of the original Dirichlet control problem has been established. Moreover, it has been shown there that the optimal solution must satisfy the first-order necessary condition associated with the optimal control problem.

Theorem 2.2. *Assume* $\epsilon \in (0, Re)$. *If* $(\mathbf{u},\mathbf{g}) \in \mathbf{W}(0,T) \times \mathcal{U}_{ad}^\epsilon$ *is an optimal solution for* $(P)_\epsilon$, *then we have*

$$\frac{1}{\epsilon}\int_0^T (\boldsymbol{\mu},\mathbf{h})_\Gamma \,\mathrm{d}t + \gamma\int_0^T (\mathbf{g},\mathbf{h})_\Gamma \,\mathrm{d}t = 0 \quad \forall \mathbf{h} \in \mathcal{U}_{ad}^\epsilon, \tag{2.7}$$

where $\boldsymbol{\mu} \in \mathbf{W}(0,T)$ *is the weak solution of the adjoint equations*

$$\begin{cases} -(\partial_t\boldsymbol{\mu},\mathbf{v}) + \frac{1}{Re}(\nabla\boldsymbol{\mu},\nabla\mathbf{v}) + c(\mathbf{u},\mathbf{v},\boldsymbol{\mu}) + c(\mathbf{v},\mathbf{u},\boldsymbol{\mu}) + \frac{1}{\epsilon}(\boldsymbol{\mu},\mathbf{v})_\Gamma \\ \qquad\qquad = (\Theta_u(\mathbf{u}),\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{V} \\ \boldsymbol{\mu}(\mathbf{x},T) = 0. \end{cases} \tag{2.8}$$

On the solutions $(\mathbf{u},p,\boldsymbol{\mu},\pi)$ of optimality system (2.6)$-$(2.8), we will make the following assumptions:

Assumption (A). There is positive integer k such that

$$\begin{cases} \sup\limits_{0<t<T}\{\|(\mathbf{u},\boldsymbol{\mu})\|_{k+1} + \|\partial_t(\mathbf{u},\boldsymbol{\mu})\|_{k-1} + \|(p,\pi)\|_k + \tau(t)\|(\partial_t\mathbf{u},\partial_t\boldsymbol{\mu})\|_k\} \le M, \\ \sup\limits_{0<t<T} \mathrm{e}^{-\alpha t}\displaystyle\int_0^t \mathrm{e}^{\alpha s}\|(\partial_t\mathbf{u},\partial_t\boldsymbol{\mu})\|_k^2 + \sigma(s)\|(\partial_t\mathbf{u},\partial_t\boldsymbol{\mu})\|_{k+1}^2 + \sigma(s)\|(\partial_t p,\partial_t\pi)\|_k^2 \,\mathrm{d}s < M, \end{cases} \tag{2.9}$$

for some positive constant M, where $\sigma(t) := \tau(t)\mathrm{e}^{\alpha t}$, $\alpha > 0$ and $\tau(t) := \min\{1,t\}$.

3. Finite element approximation of penalized control problem

In this section, we study the mixed finite element approximation of the problem in $(P)_\epsilon$. In order to fix the problem, we will set $\Theta = \Theta_1$ in the rest of this section.

3.1. Finite element spaces and properties of projections onto finite element spaces

We consider conforming mixed finite element approximations for spatial discretizations. Let $\mathcal{T}_h$ be a family of subdivisions (*e.g.* triangulation) of $\overline{\Omega} \subset \mathbb{R}^d$ satisfying $\overline{\Omega} = \cup_{K\in\mathcal{T}_h} K$ so that *diameter*$(K) \le h$ and any two closed elements K_1 and $K_2 \in \mathcal{T}_h$ are either disjoint or share exactly one face, side or vertex. Suppose further that $\mathcal{T}_h$ is a shape regular and quasi-uniform triangulation. That is, there exists a constant $C > 0$ such that the ratio between the diameter h_K of an element $K \in \mathcal{T}_h$ and the diameter of the largest ball contained in K is bounded uniformly by C, and h_K is comparable with the mesh size $h = \max\limits_{K\in\mathcal{T}_h} h_K$ for all $K \in \mathcal{T}_h$. For example, $\mathcal{T}_h$ consists of triangles that are non-degenerate as $h \to 0$. We consider conforming mixed finite element approximations for spatial discretizations. Let $\mathbf{X}_h \subset \mathbf{H}^1(\Omega)$ and $Q_h \subset L^2(\Omega)$ be a family of finite dimensional subspaces parameterized by a parameter h such that $0 < h < 1$. Here we may choose any pair of subspaces that can be used for finding finite element approximations of solutions of Navier−Stokes equations.

We make the following assumptions on the finite dimensional subspaces:

Assumption (B).
We have the approximation properties: there exists an integer k and a constant C, independent of h, $\mathbf{v}$ and q, such that

$$\inf_{\mathbf{v}_h \in \mathbf{X}_h} [\|\mathbf{v} - \mathbf{v}_h\| + h\|\nabla(\mathbf{v} - \mathbf{v}_h)\|] \leq Ch^{\ell+1}\|\mathbf{v}\|_{\ell+1} \quad \forall \mathbf{v} \in \mathbf{H}^{\ell+1}(\Omega), \quad 1 \leq \ell \leq k$$

and

$$\inf_{q_h \in Q_h} \|q - q_h\| \leq Ch^{\ell}\|q\|_{\ell} \quad \forall q \in H^{\ell}(\Omega).$$

Assumption (C). (Discrete inf-sup condition)
For every $q_h \in Q_h$, there exists a nonzero function $\mathbf{v}_h \in \mathbf{X}_h$ and $\beta > 0$ such that

$$|b(q_h, \mathbf{v}_h)| \geq \beta\|\nabla \mathbf{v}_h\|\|q_h\|,$$

with an inf-sup constant $\beta > 0$ that is independent of the mesh size h, where the bilinear form $b(\cdot, \cdot)$ is defined as $b(\mathbf{v}_h, q_h) := -\int_\Omega q_h \nabla \cdot \mathbf{v}_h \; \mathrm{d}\Omega$. The discrete inf-sup condition is needed in finite element approximations of the Navier−Stokes equations (see, *e.g.*, [10]) and naturally is also needed in the approximations of the optimality system of equation discussed below.

Assumption (D). For any integers l and m $(0 \leq l \leq m \leq 1)$ and any real numbers p and q $(1 \leq p \leq q \leq \infty)$ it holds that

$$\|\psi_h\|_{m,q} \leq ch^{l-m+d(1/q-1/p)}\|\psi_h\|_{l,p} \quad \forall \psi_h \in \mathbf{X}_h.$$

There are many conforming finite element spaces satisfying the Assumptions (B)−(D). One may choose, for example as in Section 4, the Taylor−Hood element pair for the velocity and pressure (i.e, piecewise quadratic polynomial for velocity and piecewise linear polynomial for pressure). Then, assumptions (B)−(D) hold with $k = 2$.

The mixed finite element Galerkin approximation of the state equation in (P_ϵ) is as follows: find $\mathbf{u}_h(t) \in \mathbf{X}_h$ and $p_h(t) \in Q_h$ such that for $t > 0$

$$\begin{cases} (\partial_t \mathbf{u}_h, \mathbf{v}_h) + \frac{1}{Re}(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + c(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) + \frac{1}{\epsilon}(\mathbf{u}_h, \mathbf{v}_h)_\Gamma \\ \qquad\qquad = \frac{1}{\epsilon}(\mathbf{g}_h, \mathbf{v}_h)_\Gamma + (\mathbf{f}, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{X}_h, \\ b(\mathbf{u}_h, q_h) = 0 \quad \forall q_h \in Q_h \end{cases} \tag{3.1}$$

and $\mathbf{u}_h(0) = \mathbf{u}_{0h}$, where $\mathbf{u}_{0h} \in \mathbf{X}_h$ is a suitable approximation of $\mathbf{u}_0 \in \mathbf{H}^1(\Omega)$.

The finite element approximation of the penalize control problem (P_ϵ) is to find $(\mathbf{u}_h, p_h) \in \mathbf{X}_h \times Q_h$ such that

$$(P_{h,\epsilon}) \qquad \min_{\mathbf{g}_h \in \mathbf{X}_h|_\Gamma} \mathcal{J}(\mathbf{u}_h, \mathbf{g}_h) = \int_0^T \left[\Theta(\mathbf{u}_h) + \frac{\gamma}{2}\int_\Gamma |\mathbf{g}_h|^2 \, \mathrm{d}\Gamma\right] \mathrm{d}t$$
$$\text{such that } (\mathbf{u}_h, p_h, \mathbf{g}_h) \text{ satisfies (3.1).}$$

Lemma 3.1. *If* $(\mathbf{u}_h, \mathbf{g}_h)$ *is the solution of* (3.1), *then* $\mathbf{u}_h \in L^4(\Omega \times (0,T))$ *and*

$$\sup_{t \in [0,T]} \|\mathbf{u}_h\|^2 + \frac{1}{Re}\|\nabla \mathbf{u}_h\|^2_{L^2(\mathbf{L}^2(\Omega))} + \frac{1}{\epsilon}\|\mathbf{u}_h\|^2_{L^2(\mathbf{L}^2(\Gamma))} \leq Re\|\mathbf{f}\|^2_{L^2(\mathbf{H}^1(\Omega)^*)} + \|\mathbf{u}_0\|^2 + \frac{1}{\epsilon}\|\mathbf{g}\|^2_{L^2(\mathbf{L}^2(\Gamma))}. \tag{3.2}$$

Moreover, if $(\mathbf{u}_h, \mathbf{g}_h)$ *is the solution of the semi-discrete optimal control problem* $(P_{h,\epsilon})$*, then there exists constants* c_1 *and* c_2 *such that*

$$\int_0^T \|\mathbf{g}_h\|^2 \,\mathrm{d}t \leq c_1 \quad \text{and} \quad \int_0^T \|\mathbf{u}_h\|_1^2 \leq c_2. \tag{3.3}$$

Proof. By setting $\mathbf{v}_h = \mathbf{u}_h$ in (3.1) and using the skew symmetry of the trilinear form, we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|^2 + \frac{1}{Re}\|\nabla \mathbf{u}_h\|^2 + \frac{1}{\epsilon}\|\mathbf{u}_h\|_{0,\Gamma}^2 = \frac{1}{\epsilon}(\mathbf{g}_h, \mathbf{u}_h)_\Gamma + (\mathbf{f}, \mathbf{u}_h), \tag{3.4}$$

First we note that the right hand side of (3.1) can be majorized using Young's inequality as follows

$$(\mathbf{f}, \mathbf{u}_h) + \frac{1}{\epsilon}(\mathbf{g}, \mathbf{u}_h)_\Gamma \leq \frac{Re}{2}\|\mathbf{f}\|_*^2 + \frac{1}{2Re}\|\nabla \mathbf{u}_h\|^2 + \frac{1}{2\epsilon}\|\mathbf{g}\|_{0,\Gamma}^2 + \frac{1}{2\epsilon}\|\mathbf{u}_h\|_{0,\Gamma}^2.$$

Employing this estimate in (3.4), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|^2 + \frac{1}{2Re}\|\nabla \mathbf{u}_h\|^2 + \frac{1}{2\epsilon}\|\mathbf{u}_h\|_{0,\Gamma}^2 \leq \frac{Re}{2}\|\mathbf{f}\|_*^2 + \frac{1}{2\epsilon}\|\mathbf{g}\|_{0,\Gamma}^2.$$

Integrating this with respect to time, we obtain the required *a priori* bound in (3.2). Furthermore, by Gagliardo−Nirenberg inequality and (3.2), we have that $\mathbf{u}_h \in L^4(\Omega \times (0,T))$.

In order to prove (3.3), let $\mathbf{g}_h$ be zero and $\mathbf{u}_h^*$ be the solution to the state equation in $(P_{h,\epsilon})$. Then by (3.2), we have

$$\mathcal{J}(0) \leq \tfrac{1}{2}\int_0^T \Theta(\mathbf{u}_h)\,\mathrm{d}t \leq \tfrac{1}{2}\int_0^T \|\mathbf{u}_h^*\|^2\,\mathrm{d}t \leq Re\|\mathbf{f}\|_{L^2(H^1(\Omega)^*)}^2 + \|\mathbf{u}_0\|^2.$$

Now if $(\mathbf{u}_h, p_h)$ is a solution of our optimal control problem then $\mathcal{J}(\mathbf{g}_h) \leq \mathcal{J}(0)$. We obtain the required inequality in (3.4) from this bound. □

If $(\mathbf{u}_h, p_h)$ is the solution of $(P_{h,\epsilon})$, then we can show as in the case of (P_ϵ) [21], there exists a adjoint state $(\boldsymbol{\mu}_h, \pi_h) \in \mathbf{X}_h \times Q_h$ such that $(\mathbf{u}_h, p_h, \boldsymbol{\mu}_h, \pi_h)$ satisfies the following optimality system consisting of forward-backward Navier−Stokes equations for the optimal velocity, pressure and adjoint fields, and an optimality condition for the control:

$$\left\{\begin{aligned}
(\partial_t \mathbf{u}_h, \mathbf{v}_h) + \tfrac{1}{Re}(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) &+ c(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) + \tfrac{1}{\epsilon}(\mathbf{u}_h, \mathbf{v}_h)_\Gamma \\
&= \tfrac{1}{\epsilon}(\mathbf{g}_h, \mathbf{v}_h)_\Gamma + (\mathbf{f}, \mathbf{v}_h) \qquad \forall \mathbf{v}_h \in \mathbf{X}_h, \\
b(\mathbf{u}_h, q_h) &= 0 \qquad \forall q_h \in Q_h, \\
\mathbf{u}_h(x,0) &= \mathbf{u}_{0h} \\
-(\partial_t \boldsymbol{\mu}_h, \mathbf{w}_h) + \tfrac{1}{Re}(\nabla \boldsymbol{\mu}_h, \nabla \mathbf{w}_h) &+ c(\mathbf{u}_h, \mathbf{w}_h, \boldsymbol{\mu}_h) + c(\mathbf{w}_h, \mathbf{u}_h, \boldsymbol{\mu}_h) + b(\mathbf{w}_h, \pi_h) \\
&+ \tfrac{1}{\epsilon}(\boldsymbol{\mu}_h, \mathbf{w}_h)_\Gamma \\
&= \delta(\nabla \mathbf{u}_h, \nabla \mathbf{w}_h) + \delta(\mathbf{u}_h, \mathbf{w}_h) \qquad \forall \mathbf{w}_h \in \mathbf{X}_h, \\
b(\boldsymbol{\mu}_h, r_h) &= 0 \qquad \forall r_h \in Q_h \\
\boldsymbol{\mu}_h(x,T) &= 0 \\
\gamma(\mathbf{g}_h, \mathbf{z}_h)_\Gamma &+ \tfrac{1}{\epsilon}(\boldsymbol{\mu}_h, \mathbf{z}_h)_\Gamma = 0 \qquad \forall \mathbf{z}_h \in \mathbf{X}_h|_\Gamma,
\end{aligned}\right. \tag{3.5}$$

for $t > 0$, where $\mathbf{u}_{0h} \in \mathbf{V}_h$ is a suitable approximation of $\mathbf{u}_0 \in \mathbf{V}$ such that $\|\mathbf{u}_0 - \mathbf{u}_{0h}\| \leq C\,h^{k+1}$.

3.2. Error estimates

In this section, we derive optimal order estimates for the errors $\mathbf{e}_u := \mathbf{u} - \mathbf{u}_h$ and $\mathbf{e}_{\boldsymbol{\mu}} := \boldsymbol{\mu} - \boldsymbol{\mu}_h$. First we dissociate the nonlinearity by introducing an intermediate solution $(\boldsymbol{\xi}_h, \boldsymbol{\zeta}_h)$. Let $(\boldsymbol{\xi}_h, \boldsymbol{\zeta}_h) \in \mathbf{V}_h \times \mathbf{V}_h$ be the finite element solution of a linearized optimality system given by

$$\begin{cases} (\partial_t \boldsymbol{\xi}_h, \mathbf{v}_h) + \frac{1}{Re}(\nabla \boldsymbol{\xi}_h, \nabla \mathbf{v}_h) + \frac{1}{\epsilon}(\boldsymbol{\xi}_h - \boldsymbol{\varsigma}_h, \mathbf{v}_h)_\Gamma = -c(\mathbf{u}, \mathbf{u}, \mathbf{v}_h) \\ \qquad + (\mathbf{f}, \mathbf{v}_h) \qquad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ -(\partial_t \boldsymbol{\zeta}_h, \mathbf{w}_h) + \frac{1}{Re}(\nabla \boldsymbol{\zeta}_h, \nabla \mathbf{w}_h) + \frac{1}{\epsilon}(\boldsymbol{\zeta}_h, \mathbf{w}_h)_\Gamma = -c(\mathbf{u}, \mathbf{w}_h, \boldsymbol{\mu}) - c(\mathbf{w}_h, \mathbf{u}, \boldsymbol{\mu}) \\ \qquad + \delta(\nabla \mathbf{u}, \nabla \mathbf{w}_h) + \delta(\mathbf{u}, \mathbf{w}_h) \qquad \forall \mathbf{w}_h \in \mathbf{V}_h, \\ \gamma(\boldsymbol{\varsigma}_h, z_h)_\Gamma + \frac{1}{\epsilon}(\boldsymbol{\zeta}_h, z_h)_\Gamma = 0 \qquad \forall z_h \in \mathbf{V}_h|_\Gamma, \\ \boldsymbol{\xi}_h(0) = \mathscr{P}_h \mathbf{u}_0, \quad \boldsymbol{\zeta}_h(T) = 0, \end{cases} \tag{3.6}$$

where

$$\mathbf{V}_h := \{\mathbf{v}_h \in \mathbf{X}_h |\, b(\mathbf{v}_h, q_h) = 0 \quad \forall q_h \in Q_h\}$$

is the discretely divergence free subspace of $\mathbf{X}_h$ and $\mathscr{P}_h : \mathbf{L}^2(\Omega) \to \mathbf{V}_h$ is the L^2-orthogonal projection. Let $\widetilde{\mathbf{e}}_u := \mathbf{u} - \boldsymbol{\xi}_h$, $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}} := \boldsymbol{\mu} - \boldsymbol{\zeta}_h$, $\widehat{\mathbf{e}}_u := \boldsymbol{\xi}_h - \mathbf{u}_h$, $\widehat{\mathbf{e}}_u := \boldsymbol{\xi}_h - \mathbf{u}_h$, $\widetilde{\mathbf{e}}_g := \mathbf{g} - \boldsymbol{\varsigma}_h$ and $\widehat{\mathbf{e}}_g := \boldsymbol{\varsigma}_h - \mathbf{g}_h$ so that $\mathbf{e}_u = \widetilde{\mathbf{e}}_u + \widehat{\mathbf{e}}_u$, $\quad \mathbf{e}_{\boldsymbol{\mu}} = \widetilde{\mathbf{e}}_{\boldsymbol{\mu}} + \widehat{\mathbf{e}}_{\boldsymbol{\mu}}$ and $\mathbf{e}_g = \widetilde{\mathbf{e}}_g + \widehat{\mathbf{e}}_g$. Below, we derive some estimates for $\widetilde{\mathbf{e}}_u$ and $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}$.

Lemma 3.2. *Suppose that the Assumptions* (A), (B) *hold and that* $\epsilon \in (0, Re)$ *and* $\epsilon = \mathcal{O}(h)$. *Let* $(\boldsymbol{\xi}_h, \boldsymbol{\zeta}_h) \in \mathbf{V}_h \times \mathbf{V}_h$ *be the solution of* A. *Then there exists a constant* C, *independent of* h, *such that the errors* $\widetilde{\mathbf{e}}_u := (\mathbf{u} - \boldsymbol{\xi}_h)$ *and* $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}} : (\boldsymbol{\mu} - \boldsymbol{\zeta}_h)$ *satisfy*

$$\int_0^t \mathrm{e}^{\alpha\, s} \|(\widetilde{\mathbf{e}}_u, \widetilde{\mathbf{e}}_{\boldsymbol{\mu}})\|^2 \,\mathrm{d}s + \frac{1}{\epsilon} \int_0^T \mathrm{e}^{\alpha\, s} \|\widetilde{\mathbf{e}}_u\|_{0,\Gamma}^2 \,\mathrm{d}s + \frac{1}{\epsilon^3} \int_0^T \mathrm{e}^{\alpha\, s} \|\widetilde{\mathbf{e}}_\mu\|_{0,\Gamma}^2 \,\mathrm{d}s \le C\, \sigma(t)\, h^{2k+2}, \tag{3.7}$$

$0 \le t < T$, *where* $\sigma(t) := \tau(t)\mathrm{e}^{\alpha t}$, $\alpha > 0$ *and* $\tau(t) := \min\{1, t\}$.

Proof. First notice that $(\widetilde{\mathbf{e}}_u, \widetilde{\mathbf{e}}_{\boldsymbol{\mu}})$ satisfies

$$\begin{cases} (\partial_t \widetilde{\mathbf{e}}_u, \mathbf{v}_h) + \frac{1}{Re}(\nabla \widetilde{\mathbf{e}}_u, \nabla \mathbf{v}_h) + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_u, \mathbf{v}_h)_\Gamma = \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathbf{v}_h)_\Gamma \\ \qquad + b(\mathbf{v}_h, p) \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ -(\partial_t \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \mathbf{w}_h) + \frac{1}{Re}(\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \nabla \mathbf{w}_h) + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \mathbf{w}_h)_\Gamma = b(\mathbf{w}_h, \pi) \quad \forall \mathbf{w}_h \in \mathbf{V}_h, \\ \gamma(\widetilde{\mathbf{e}}_g, \mathbf{z}_h)_\Gamma + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \mathbf{z}_h)_\Gamma = 0 \quad \mathbf{z}_h \in \mathbf{V}_h|_\Gamma, \\ \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}(T) = 0 \quad \text{and} \quad \widetilde{\mathbf{e}}_u(0) = \mathbf{u}_0 - \mathscr{P}_h \mathbf{u}_0. \end{cases} \tag{3.8}$$

Setting $\mathbf{w}_h = \mathscr{P}_h \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}$ in $(3.8)_2$ and using the approximation properties of $\mathscr{P}_h$ yields

$$\begin{aligned} -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2 + \frac{1}{Re}\|\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2 + \frac{1}{\epsilon}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|_{0,\Gamma}^2 &\le \frac{1}{Re}(\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \nabla(\boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu})) \\ &\quad + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}, \boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu})_\Gamma - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu}\|^2 - (\nabla \cdot (\mathscr{P}_h \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}), \pi) \\ &\le ch^k \|\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\| \|\boldsymbol{\mu}\|_{k+1} + \frac{1}{\epsilon}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|_{0,\Gamma} \|\boldsymbol{\mu} - \mathscr{P}_h \mu\|_{0,\Gamma} \\ &\quad + ch^k \|\pi\|_k \|\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\| - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu}\|^2. \end{aligned}$$

Now by Young's inequality, we have

$$-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2 + \frac{1}{2Re}\|\nabla \widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2 + \frac{1}{2\epsilon}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|_{0,\Gamma}^2 \le ch^{2k}(\|\boldsymbol{\mu}\|_{k+1}^2 + \|\pi\|_k^2) + \frac{1}{\epsilon}\|\boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu}\|_{0,\Gamma}^2 - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\mu} - \mathscr{P}_h \boldsymbol{\mu}\|^2.$$

Multiplying by $\mathrm{e}^{\alpha t}$, we find

$$-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2)+\alpha\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2+\frac{\mathrm{e}^{\alpha t}}{2Re}\|\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2+\frac{\mathrm{e}^{\alpha t}}{2\epsilon}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\le ch^{2k}\mathrm{e}^{\alpha t}(\|\boldsymbol{\mu}\|^2_{k+1}+\|\pi\|^2_k)$$
$$+\frac{\mathrm{e}^{\alpha t}}{\epsilon}\|\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu}\|^2_{0,\Gamma}-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\alpha t}\|\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu}\|^2)+\frac{\alpha}{2}\mathrm{e}^{\alpha t}\|\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu}\|^2.$$

Using the estimate $\|\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu}\|_{0,\Gamma}\le ch^{1/2}\|\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu}\|_1$ (see [12]), and integrating with respect to time yields

$$\frac{1}{Re}\int_t^T \mathrm{e}^{\alpha s}\|\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s+\frac{1}{\epsilon}\int_t^T \mathrm{e}^{\alpha s}\|\mathbf{e}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s\le C\,h^{2k}(1+\frac{h}{\epsilon})\int_t^T \mathrm{e}^{\alpha s}(\|\boldsymbol{\mu}\|^2_{k+1}+\|\pi\|^2_k)\,\mathrm{d}s. \tag{3.9}$$

Let us next employ a duality argument to derive optimal order error estimates in L^2-norm. Let $(\boldsymbol{\Phi},r)\in\mathbf{V}\times L^2_0(\Omega)$ be the unique solution to the problem

$$\begin{cases}(\partial_s\Phi,\psi)+\frac{1}{Re}(\nabla\Phi,\nabla\psi)+(\nabla\cdot\psi,r)+\frac{1}{\epsilon}(\Phi,\psi)_\Gamma & \\ \qquad\qquad =(\mathrm{e}^{\alpha s}\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\psi)+\frac{1}{\epsilon^3}(\mathrm{e}^{\alpha s}\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\psi)_\Gamma \quad \psi\in\mathbf{H}^1(\Omega), \\ \Phi(0) \quad =0,\quad 0\le s\le t.\end{cases} \tag{3.10}$$

It can be easily shown that the following *a priori* estimates hold

$$\int_0^T \mathrm{e}^{-\alpha s}\left\{\|\partial_s\Phi\|^2+\|\Phi\|^2_{k+1}+\|r\|^2_k\right\}\,\mathrm{d}s\le c\left[\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s+\frac{1}{\epsilon^3}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s\right]. \tag{3.11}$$

Setting $\psi=\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}$ in (3.10) and $\mathbf{w}_h=\mathscr{P}_h\Phi$ in $(3.8)_2$ yield, respectively,

$$\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2+\frac{\mathrm{e}^{\alpha t}}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}=(\partial_\tau\Phi,\widetilde{\mathbf{e}}_{\boldsymbol{\mu}})+\frac{1}{Re}(\nabla\Phi,\nabla\widetilde{\mathbf{e}}_\mu)+(\nabla\cdot\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},r)+\frac{1}{\epsilon}(\Phi,\widetilde{\mathbf{e}}_{\boldsymbol{\mu}})_\Gamma \tag{3.12}$$

and

$$(\partial_\tau\Phi,\widetilde{\mathbf{e}}_{\boldsymbol{\mu}})=\frac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\Phi)+(\partial_\tau\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi-\Phi)-\frac{1}{Re}(\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\nabla\mathscr{P}_h\Phi)-\frac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi)_\Gamma+(\nabla\cdot\mathscr{P}_h\Phi,\pi). \tag{3.13}$$

Employing (3.13) in the right-hand side of (3.12), we obtain

$$\begin{aligned}\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2+\tfrac{\mathrm{e}^{\alpha t}}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}&=\tfrac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\Phi)-(\partial_\tau\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi-\Phi)+\tfrac{1}{Re}(\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\nabla(\Phi-\mathscr{P}_h\Phi))\\&\quad+(\nabla\cdot\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},r)+(\nabla\cdot\mathscr{P}_h\Phi,\pi)+\tfrac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\Phi-\mathscr{P}_h\Phi)_\Gamma.\end{aligned} \tag{3.14}$$

From the definition of $\mathscr{P}_h$, the second term on the right-hand side becomes

$$(\partial_\tau\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi-\Phi)=\frac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi-\Phi)+(\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu},\partial_\tau\Phi). \tag{3.15}$$

Using (3.15) in (3.14), we obtain

$$\begin{aligned}\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2+\frac{\mathrm{e}^{\alpha t}}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}&=\frac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\mathscr{P}_h\Phi)+(\partial_\tau\Phi,\boldsymbol{\mu}-\mathscr{P}_h\boldsymbol{\mu})+\frac{1}{Re}(\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\nabla(\Phi-\mathscr{P}_h\Phi))\\&\quad+(\nabla\cdot\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},r-j_hr)+(\nabla\cdot(\mathscr{P}_h\Phi-\Phi),\pi-j_h\pi)\\&\quad+\frac{1}{\epsilon}(\widetilde{\mathbf{e}}_{\boldsymbol{\mu}},\Phi-\mathscr{P}_h\Phi)_\Gamma,\end{aligned} \tag{3.16}$$

for $j_h\pi, j_h r \in Q_h$. Integrating (3.16) with respect to τ and noting that $\Phi(0) = 0$ and $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}(T) = 0$, we obtain by the approximation properties that

$$\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s + \frac{1}{\epsilon^3}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s \le ch^{2k+2}\int_0^T \mathrm{e}^{\alpha s}(\|\boldsymbol{\mu}\|^2_{k+1} + \|\pi\|^2_k)\,\mathrm{d}s$$
$$+ ch^{2k}\int_0^T \mathrm{e}^{\alpha s}\|\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s + c\frac{h^{2k+1}}{\epsilon^2}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s$$
$$+ \widehat{\epsilon}\int_0^T \mathrm{e}^{-\alpha s}(\|\partial_t\Phi\|^2 + \|\Phi\|^2_{k+1} + \|r\|^2_k)\,\mathrm{d}s. \tag{3.17}$$

Using (3.11) in (3.17) and choosing $\widehat{\epsilon}$ sufficiently small, we obtain

$$\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s + \frac{1}{\epsilon^3}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s \le ch^{2k+2}\int_0^T \mathrm{e}^{\alpha s}\left(\|\boldsymbol{\mu}\|^2_{k+1} + \|\pi\|^2_k\right)\,\mathrm{d}s$$
$$+ c\left(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2}\right)\left\{h^{2k}\left(\int_0^T \mathrm{e}^{\alpha s}\|\nabla\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s + \frac{1}{\epsilon}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s\right\}\right.$$

for some constant c.

This combined with (3.9) yields

$$\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2\,\mathrm{d}s + \frac{1}{\epsilon^3}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}\,\mathrm{d}s \le c\left(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2}\right)\sigma(t)h^{2k+2}. \tag{3.18}$$

In order to obtain an analogous estimate for $\widetilde{\mathbf{e}}_u$, we proceed along similar lines. First by setting $\mathbf{w}_h = \mathscr{P}_h\widetilde{\mathbf{e}}_u$ in $(3.8)_1$, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{Re}\|\nabla\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma} \le \frac{1}{Re}(\nabla\widetilde{\mathbf{e}}_u, \nabla(\mathbf{u} - \mathscr{P}_h\mathbf{u})) + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_u, \mathbf{u} - \mathscr{P}_h\mathbf{u})_\Gamma$$
$$+ \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u} - \mathscr{P}_h\mathbf{u}\|^2 - (\nabla\cdot(\mathscr{P}_h\widetilde{\mathbf{e}}_u), p) + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \widetilde{\mathbf{e}}_u)_\Gamma + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathscr{P}_h\mathbf{u} - \mathbf{u})_\Gamma. \tag{3.19}$$

Let us start estimating the terms involving $\widetilde{\mathbf{e}}_g$. First notice by Young's inequality

$$\frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \widetilde{\mathbf{e}}_u)_\Gamma + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathscr{P}_h\mathbf{u} - \mathbf{u})_\Gamma \le \frac{3}{2\epsilon}\|\widetilde{\mathbf{e}}_g\|^2_{0,\Gamma} + \frac{1}{4\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma} + \frac{1}{2\epsilon}\|\mathscr{P}_h\mathbf{u} - \mathbf{u}\|^2_{0,\Gamma}. \tag{3.20}$$

Now by putting $z_h = \mathscr{P}_h(\widetilde{\mathbf{e}}_g)$ in $(3.8)_3$ and rearranging the terms, we obtain

$$\gamma\|\widetilde{\mathbf{e}}_g\|^2_{0,\Gamma} = \gamma(\widetilde{\mathbf{e}}_g, \mathbf{g} - \mathscr{P}_h\mathbf{g})_\Gamma + \frac{1}{\epsilon}(\boldsymbol{\mu} - \boldsymbol{\zeta}_h, \mathbf{g} - \widetilde{\mathbf{g}}_h)_\Gamma + \frac{1}{\epsilon}(\boldsymbol{\mu} - \boldsymbol{\zeta}_h, \mathscr{P}_h\mathbf{g} - \mathbf{g})_\Gamma.$$

Therefore by Cauchy−Schwarz inequality, we have

$$\|\widetilde{\mathbf{e}}_g\|_{0,\Gamma} \le c\|\mathbf{g} - \mathscr{P}_h\mathbf{g}\|_{0,\Gamma} + \frac{c}{\gamma\epsilon}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|_{0,\Gamma}. \tag{3.21}$$

Thus by (3.20) and (3.21), we have

$$\frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \widetilde{\mathbf{e}}_u)_\Gamma + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathscr{P}_h\mathbf{u} - \mathbf{u})_\Gamma \le \frac{c}{\epsilon}\|\mathbf{g} - \mathscr{P}_h\mathbf{g}\|^2_{0,\Gamma} + \frac{C}{\gamma^2\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma} + \frac{1}{4\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}$$
$$+ \frac{1}{2\epsilon}\|\mathbf{u} - \mathscr{P}_h\mathbf{u}\|^2_{0,\Gamma}. \tag{3.22}$$

Employing (3.22) in (3.19) and estimating other terms as before, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{Re}\|\nabla\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{2\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma} \leq c\delta^{-1}h^{2k}(\|\mathbf{u}\|^2_{k+1} + \|p\|^2_k) + \delta\|\nabla\widetilde{\mathbf{e}}_u\|^2 + \frac{c}{\epsilon}\|\mathbf{u} - \mathscr{P}_h\mathbf{u}\|^2_{0,\Gamma}$$
$$+ \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u} - \mathscr{P}_h\mathbf{u}\|^2 + \frac{c}{\epsilon}\|\mathbf{g} - \mathscr{P}_h\mathbf{g}\|^2_{0,\Gamma} + \frac{c}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}.$$

Multiplying by $\mathrm{e}^{\alpha t}$ and using Poincaré−Friedrich's inequality, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\alpha t}\|\widetilde{\mathbf{e}}_u\|^2) + \mathrm{e}^{\alpha t}\left(1 - \frac{\alpha\, Re}{\lambda} - \delta\right)\left(\frac{1}{Re}\|\nabla\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{2\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\right) \leq c\delta^{-1}\mathrm{e}^{\alpha\, t}h^{2k}(\|\mathbf{u}\|^2_{k+1} + \|p\|^2_k)$$
$$+ \frac{c}{\epsilon}h^{2k}\mathrm{e}^{\alpha t}\|\mathbf{u}\|_{k+1} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\alpha t}\|\mathbf{u} - \mathscr{P}_h\mathbf{u}\|^2) + \frac{c\mathrm{e}^{\alpha t}}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|^2_{0,\Gamma}.$$

We can now choose $\delta = \frac{1}{2} - \frac{\alpha\, Re}{2\lambda}$ and integrate with respect to time t and use (3.18), to obtain

$$\int_0^T \mathrm{e}^{\alpha s}\|\nabla\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}s + \frac{1}{\epsilon}\int_0^T \mathrm{e}^{\alpha\, s}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\,\mathrm{d}s \leq c\left(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2}\right)h^{2k}\left(1 - \frac{\alpha\, Re}{\lambda}\right)^{-2}\int_0^t \mathrm{e}^{\alpha s}(\|\mathbf{u}\|^2_{k+1} + \|p\|^2_k)\,\mathrm{d}s. \tag{3.23}$$

In order to derive the optimal order error in L^2-norm, we use the following duality argument. For fixed $h > 0$ and $t > 0$, let $(\varXi(\tau), s(\tau))$ be the unique solution of the backward problem:

$$\begin{cases} (\partial_\tau\varXi, \mathbf{w}) - \frac{1}{Re}(\nabla\varXi, \nabla\mathbf{w}) - (\nabla\cdot\mathbf{w}, \hat{p}) - \frac{1}{\epsilon}(\varXi, \mathbf{w})_\Gamma \\ \qquad = (\mathrm{e}^{\alpha s}\,\widetilde{\mathbf{e}}_u, \mathbf{w}) + \frac{1}{\epsilon}(\mathrm{e}^{\alpha s}\widetilde{\mathbf{e}}_u, \mathbf{w})_\Gamma \quad \forall\mathbf{w}\in\mathbf{H}^1(\Omega) \\ \varXi(t) = 0, \quad 0\leq\tau\leq t. \end{cases} \tag{3.24}$$

With a change of variable $t \to t - \tau$, set $\varXi(\tau) = \varXi(t - \tau)$. Then $\varXi(\tau)$ satisfies a forward linear unsteady Stokes type problem. Thus, we obtain the following *a priori* estimate

$$\int_0^t \mathrm{e}^{-\alpha s}\left(\|\partial_t\varXi\|^2 + \|\varXi\|^2_{k+1} + \|\hat{p}\|^2_k + \frac{1}{\epsilon}\|\varXi\|^2_{k+1/2,\Gamma}\right)\mathrm{d}s \leq C\left(1 - \frac{\alpha\, Re}{\lambda}\right)^{-2}\left[\int_0^t \mathrm{e}^{\alpha s}(\|\widetilde{\mathbf{e}}_u\|^2 + \frac{1}{\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma})\,\mathrm{d}s\right]. \tag{3.25}$$

Setting $\mathbf{w} = \widetilde{\mathbf{e}}_u$ in (3.24) and $\mathbf{v}_h = \mathscr{P}_h\varXi$ in $(3.8)_1$ yield, respectively,

$$\mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2 + \frac{\mathrm{e}^{\alpha s}}{\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma} = (\partial_s\varXi, \widetilde{\mathbf{e}}_u) - \frac{1}{Re}(\nabla\varXi, \nabla\widetilde{\mathbf{e}}_u) - (\nabla\cdot\widetilde{\mathbf{e}}_u, \hat{p}) - \frac{1}{\epsilon}(\varXi, \widetilde{\mathbf{e}}_u)_\Gamma \tag{3.26}$$

and

$$(\partial_\tau\varXi, \widetilde{\mathbf{e}}_u) = \frac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_u, \varXi) + (\partial_\tau\widetilde{\mathbf{e}}_u, \mathscr{P}_h\varXi - \varXi) + \frac{1}{Re}(\nabla\widetilde{\mathbf{e}}_u, \nabla\mathscr{P}_h\varXi) - \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathscr{P}_h\varXi)_\Gamma + \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_u, \mathscr{P}_h\varXi)_\Gamma$$
$$- (\nabla\cdot\mathscr{P}_h\varXi, p). \tag{3.27}$$

Using (3.27) in the right-hand side of (3.26) and using the definition of $\mathscr{P}_h$, we obtain

$$\mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2 + \frac{\mathrm{e}^{\alpha s}}{\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma} = \frac{\mathrm{d}}{\mathrm{d}\tau}(\widetilde{\mathbf{e}}_u, \mathscr{P}_h\varXi) + (\partial_\tau\varXi, \mathbf{u} - \mathscr{P}_h\mathbf{u}) + \frac{1}{Re}(\nabla\widetilde{\mathbf{e}}_u, \nabla(\mathscr{P}_h\varXi - \varXi))$$
$$- (\nabla\cdot\widetilde{\mathbf{e}}_u, \hat{p} - j_h\hat{p}) - (\nabla\cdot(\mathscr{P}_h\varXi - \varXi), p - j_hp)$$
$$+ \frac{1}{\epsilon}(\widetilde{\mathbf{e}}_u, \mathscr{P}_h\varXi - \varXi)_\Gamma - \frac{1}{\epsilon}\left(\mathscr{P}_h\varXi, \frac{1}{\gamma\epsilon}\widetilde{\mathbf{e}}_\mu\right)_\Gamma. \tag{3.28}$$

Integrating (3.28) with respect to τ and using the fact that $\Xi(t)=0$ and $\widetilde{\mathbf{e}}_u(0)=\mathbf{u}_0-\mathscr{P}_h\mathbf{u}_0$, we obtain

$$\begin{aligned}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}\tau+\frac{1}{\epsilon}\int_0^T \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\,\mathrm{d}s &\le ch^{2k}\int_0^t \mathrm{e}^{\alpha s}\|\nabla\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}\tau\\ &+ch^{2k+2}\int_0^t \mathrm{e}^{\alpha s}(\|\mathbf{u}\|^2_{k+1}+\|p\|^2_k)\,\mathrm{d}s+\frac{h^{2k+1}}{\epsilon^2}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\,\mathrm{d}s+\frac{C}{\epsilon^3}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_\mu\|^2_{0,\Gamma}\,\mathrm{d}s\\ &+\widehat{\epsilon}\left[\int_0^t \mathrm{e}^{\alpha s}(\|\partial_t\Xi\|^2+\|\Xi\|^2_{k+1}+\|\hat{p}\|^2_k+\frac{1}{\epsilon}\|\Xi\|^2_{0,\Gamma})\,\mathrm{d}s\right]. \end{aligned}\tag{3.29}$$

Using (3.25) in (3.29) and choosing $\widehat{\epsilon}$ sufficiently small, we obtain

$$\begin{aligned}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}\tau+\frac{1}{\epsilon}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\,\mathrm{d}\tau &\le c(1-\frac{\alpha\,Re}{\lambda})^{-2}(1+\frac{h}{\epsilon}+\frac{h^2}{\epsilon^2}+\frac{h^3}{\epsilon^3})\\ &\Big\{h^{2k+2}\int_0^t \mathrm{e}^{\alpha s}((\|\mathbf{u}\|^2_{k+1}+\|p\|^2_k)\,\mathrm{d}s\\ &+h^{2k}\int_0^t \mathrm{e}^{\alpha s}(\|\widetilde{\mathbf{e}}_u\|^2+\frac{1}{\epsilon}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma})\,\mathrm{d}s+\frac{1}{\epsilon^3}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_\mu\|^2_{0,\Gamma}\,\mathrm{d}s\Big\}.\end{aligned}$$

This combined with (3.23) yields

$$\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}\tau+\frac{1}{\epsilon}\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_u\|^2_{0,\Gamma}\,\mathrm{d}\tau\le c\left(1+\frac{h}{\epsilon}+\frac{h^2}{\epsilon^2}+\frac{h^3}{\epsilon^3}\right)\sigma(t)h^{2k+2}. \tag{3.30}$$

Combining (3.18) and (3.30) and choosing $\epsilon=\mathcal{O}(h)$, we obtain (3.7). □

For optimal order error estimates of $(\widetilde{\mathbf{e}}_u,\widetilde{\mathbf{e}}_{\boldsymbol{\mu}})$ in the $L^\infty(\mathbf{L}^2(\Omega))$ and $L^\infty(\mathbf{H}^1(\Omega))$-norms, we introduce the Stokes projection $\mathscr{P}_h^s:\mathbf{H}^1(\Omega)\to\mathbf{V}_h$ according to the rule

$$\frac{1}{Re}(\nabla(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon}),\nabla\mathbf{v}_h)+\frac{1}{\epsilon}(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon},\mathbf{v}_h)_\Gamma=(\nabla\cdot\mathbf{v}_h,\iota)\quad\forall\mathbf{v}_h\in\mathbf{V}_h. \tag{3.31}$$

With Stokes projection defined as above, we now decompose $(\widetilde{\mathbf{e}}_u,\widetilde{\mathbf{e}}_{\boldsymbol{\mu}})$ as

$$\widetilde{\mathbf{e}}_u:=(\mathbf{u}-\mathscr{P}_h^s\mathbf{u})+(\mathscr{P}_h^s\mathbf{u}-\boldsymbol{\xi}_h)\quad\text{and}\quad\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}:=(\boldsymbol{\mu}-\mathscr{P}_h^s\boldsymbol{\mu})+(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\zeta}_h).$$

First of all, we derive optimal error estimates for the error in Stokes projection.

Lemma 3.3. *Suppose that Assumptions* (A) *and* (B) *hold and* $\epsilon=\mathcal{O}(h)$. *Then there exists a constant* C, *independent of* h, *such that the error in Stokes projection satisfies the following*

(i)

$$\|\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon}\|+h\|\nabla(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon})\|+\frac{1}{\epsilon}\|\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}\le C\,h^{k+1}(\|\boldsymbol{\Upsilon}\|_{k+1}+\|\iota\|_k),$$

(ii)

$$\begin{aligned}&\|\partial_t(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon})\|+h\|\partial_t(\nabla(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon})\|\,\mathrm{d}s+\tfrac{1}{\epsilon}\|\partial_t(\boldsymbol{\Upsilon}-\mathscr{P}_h^s\boldsymbol{\Upsilon})\|_{0,\Gamma}\\ &\le C\,h^{k+1}(\|\partial_t\boldsymbol{\Upsilon}\|_{k+1}+\|\partial_t\iota\|_k),\end{aligned}$$

for all $t\in[0,T)$.

Proof.

(i) Setting $\mathbf{v}_h = \mathscr{P}_h(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})$ in (3.31) yields

$$\frac{1}{Re}\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|^2 + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}^2 \le c\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\| + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}\|\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon}\|_{0,\Gamma} + c\|\iota - j_h\iota\|\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|.$$

Using the estimate $\|\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon}\|_{0,\Gamma} \le ch^{1/2}\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|$ (see [12]) and the approximation properties of the L^2-projections (see [10]) we are led to

$$\frac{1}{Re}\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|^2 + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}^2 \le c\left(1 + \frac{h}{\epsilon}\right)h^{2k}. \tag{3.32}$$

For the L^2-norm estimate, we use a duality argument. Suppose now that the pair $(\widehat{\boldsymbol{\Upsilon}}, \hat{\iota})$ is the unique solution to

$$\begin{cases} \frac{1}{Re}(\nabla\widehat{\boldsymbol{\Upsilon}}, \nabla\Psi) + \frac{1}{\epsilon}(\widehat{\boldsymbol{\Upsilon}}, \Psi)_\Gamma = (\nabla\cdot\Psi, \hat{\iota}) + (\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}, \Psi) \\ \qquad\qquad + \frac{1}{\epsilon}(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}, \Psi)_\Gamma \quad \forall\Psi \in \mathbf{H}^1(\Omega) \\ (\nabla\cdot\widehat{\boldsymbol{\Upsilon}}, q) = 0 \quad \forall q \in L_0^2(\Omega). \end{cases} \tag{3.33}$$

It can be shown by standard energy arguments that

$$\|\widehat{\boldsymbol{\Upsilon}}\|_{k+1} + \|\hat{\iota}\|_k + \frac{1}{\sqrt{\epsilon}}\|\widehat{\boldsymbol{\Upsilon}}\|_{k+1/2} \le c\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\| + \frac{1}{\sqrt{\epsilon}}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}. \tag{3.34}$$

Setting $\Psi = \boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}$ in (3.33) and using the definition of Stokes projection yields

$$\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|^2 + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}^2 \le \frac{1}{Re}(\nabla(\widehat{\boldsymbol{\Upsilon}} - \mathscr{P}_h\widehat{\boldsymbol{\Upsilon}}), \nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})) + \frac{1}{\epsilon}(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}, \widehat{\boldsymbol{\Upsilon}} - \mathscr{P}_h\widehat{\boldsymbol{\Upsilon}})_\Gamma + (\nabla\cdot(\mathscr{P}_h\widehat{\boldsymbol{\Upsilon}} - \widehat{\boldsymbol{\Upsilon}}), p - j_hp)_\Gamma - (\nabla\cdot(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon}), \hat{\iota} - j_h\hat{\iota}) - (\nabla\cdot\mathscr{P}_h(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}), \iota - j_h\iota).$$

Estimating the terms on the right-hand side yields

$$\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|^2 + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}^2 \le \frac{1}{Re}\|\nabla(\widehat{\boldsymbol{\Upsilon}} - \mathscr{P}_h\widehat{\boldsymbol{\Upsilon}})\|\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\| + \frac{1}{\epsilon}\|\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon}\|_{0,\Gamma}\|\widehat{\boldsymbol{\Upsilon}} - \mathscr{P}_h\widehat{\boldsymbol{\Upsilon}}\|_{0,\Gamma} + c\|\nabla(\mathscr{P}_h\widehat{\boldsymbol{\Upsilon}} - \widehat{\boldsymbol{\Upsilon}})\|\|\iota - j_h\iota\| + c\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|\|\hat{\iota} - j_h\hat{\iota}\| + \|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|\|\iota - j_h\iota\|.$$

Applying Young inequality and the approximation properties of $\mathscr{P}_h$, we obtain the desired estimate.

(ii) By differentiating (3.31) with respect to t and setting $\mathbf{v}_h = \mathscr{P}_h\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})$, we obtain

$$\frac{1}{Re}\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|^2 + \frac{1}{\epsilon}\|\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|_{0,\Gamma}^2 \le c\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\| + \frac{1}{\epsilon}\|\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|_{0,\Gamma}\|\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|_{0,\Gamma} + c\|\iota_t - j_h\iota_t\|\|\nabla\cdot\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h^s\boldsymbol{\Upsilon})\|.$$

Using the estimate $\|\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon}\|_{0,\Gamma} \le ch^{1/2}\|\nabla(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|$ (see [12]) and the approximation properties of the projections, we are led to

$$\begin{aligned}\frac{1}{Re}\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}^s_h\boldsymbol{\Upsilon})\|^2 + \frac{1}{\epsilon}\|\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}^s_h\boldsymbol{\Upsilon})\|^2_{0,\Gamma} &\le c\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|^2 + \frac{ch}{\epsilon}\|\nabla\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}_h\boldsymbol{\Upsilon})\|^2 \\ &\quad + c\|\iota_t - j_h\iota_t\|^2 \\ &\le c(1+\frac{h}{\epsilon})h^{2k}(\|\partial_t\boldsymbol{\Upsilon}\|^2_{k+1} + \|\partial_t\iota\|^2_k).\end{aligned}$$

The result now follows by assuming $\epsilon = \mathcal{O}(h)$. In order to derive a bound for $\|\partial_t(\boldsymbol{\Upsilon} - \mathscr{P}^s_h\boldsymbol{\Upsilon})\|$, we can use duality argument as in the proof of part (i) and thus the details are skipped. □

Now we begin estimating $\widetilde{\mathbf{e}}_u := \mathbf{u} - \boldsymbol{\xi}_h$ and $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}} := \boldsymbol{\mu} - \boldsymbol{\zeta}_h$ in the $L^\infty(\mathbf{L}^2)$ and $L^\infty(\mathbf{H}^1)$-norms. Since $\widetilde{\mathbf{e}}_u := (\mathbf{u} - \mathscr{P}^s_h\mathbf{u}) + (\mathscr{P}^s_h\mathbf{u} - \boldsymbol{\xi}_h)$ and $\widetilde{\mathbf{e}}_{\boldsymbol{\mu}} := (\boldsymbol{\mu} - \mathscr{P}^s_h\boldsymbol{\mu}) + (\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)$ and the estimates of $\mathbf{u} - \mathscr{P}^s_h\mathbf{u}$ and $\boldsymbol{\mu} - \mathscr{P}^s_h\boldsymbol{\mu}$ are known from Lemma 3.3, it is sufficient to estimate $\mathscr{P}^s_h\mathbf{u} - \boldsymbol{\xi}_h$ and $\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h$.

Lemma 3.4. *Suppose that the Assumptions* (A) *and* (B) *and* (D) *hold and* $\epsilon = \mathcal{O}(h)$. *Then there exists a constant* C*, independent of* h*, such that the following estimate holds for* $(\mathscr{P}^s_h\mathbf{u} - \boldsymbol{\xi}_h, \mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)$*:*

$$\|\mathscr{P}^s_h\mathbf{u} - \boldsymbol{\xi}_h\|^2 + h^2\|\nabla(\mathscr{P}^s_h\mathbf{u} - \boldsymbol{\xi}_h)\|^2 \le Ch^{2k+2} \tag{3.35}$$

and

$$\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2 + h^2\|\nabla(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)\|^2 \le Ch^{2k+2}. \tag{3.36}$$

Proof. Let us first prove (3.36). Notice that we can write $(3.8)_2$ as

$$\begin{aligned}(\partial_t(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h), \mathbf{w}_h) + \frac{1}{Re}(\nabla(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h), \nabla\mathbf{w}_h) + \frac{1}{\epsilon}(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h, \mathbf{w}_h)_\Gamma \\ = (\partial_t(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\mu}), \mathbf{w}_h) \quad \mathbf{w}_h \in \mathbf{V}_h.\end{aligned} \tag{3.37}$$

Setting $\mathbf{w}_h = \sigma(t)(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)$ in (3.37) yields

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\sigma(t)\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2) + \frac{\sigma(t)}{Re}\|\nabla(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)\|^2 + \frac{\sigma(t)}{\epsilon}\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2_{0,\Gamma} \\ \le \sigma(t)\|\partial_t(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\mu})\|\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\| + \frac{\sigma_t}{2}\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2\end{aligned} \tag{3.38}$$

where $\sigma(t) := \tau(t)\mathrm{e}^{\alpha t}$, $\alpha > 0$ and $\tau(t) := \min\{1, t\}$. Therefore applying Young's inequality and integrating with respect to time t yields

$$\begin{aligned}\sigma(t)\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2 + \int_0^t \frac{\sigma(t)}{Re}\|\nabla(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h)\|^2\,\mathrm{d}s + \int_0^t \frac{\sigma(t)}{\epsilon}\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2_{0,\Gamma}\,\mathrm{d}s \\ \le c\int_0^t \frac{\sigma^2}{\partial_s\sigma}\|\partial_s(\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\mu})\|^2\,\mathrm{d}s + c\int_0^t \mathrm{e}^{\alpha s}\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\mu}\|^2\,\mathrm{d}s \\ + c\int_0^t \mathrm{e}^{\alpha s}\|\widetilde{\mathbf{e}}_\mu\|^2\,\mathrm{d}s.\end{aligned} \tag{3.39}$$

Since $\frac{\sigma^2}{\partial_s\sigma} \le \frac{\tau(s)\sigma(s)}{\alpha}$, using the estimate in Lemma 3.3 for $\|\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|$ and Lemma 3.2, we obtain

$$\begin{aligned}\|\mathscr{P}^s_h\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2 &\le ch^{2k+2}\sigma^{-1}(t)\int_0^t \tau(s)\sigma(s)(\|\boldsymbol{\mu}_s\|^2_{k+1} + \|\pi_s\|^2_k)\mathrm{d}s \\ &\quad + ch^{2k+2}\sigma^{-1}(t)\int_0^t \mathrm{e}^{\alpha s}(\|\boldsymbol{\mu}\|^2_{k+1} + \|\pi\|^2_k)\,\mathrm{d}s + ch^{2k+2}.\end{aligned}$$

In order to bound the first term on the right-hand side of last inequality, we note that for $0 < t \leq 1$,

$$\sigma^{-1}(t)\int_0^t \tau(s)\sigma(s)(\|\boldsymbol{\mu}_s\|_{k+1}^2 + \|\pi_s\|_k^2)\mathrm{d}s \leq t^{-1}\mathrm{e}^{-\alpha\, t}\int_0^t s\sigma(s)(\|\boldsymbol{\mu}_s\|_{k+1}^2 + \|\pi_s\|_k^2)\mathrm{d}s$$
$$\leq \mathrm{e}^{-\alpha\, t}\int_0^t \sigma(s)(\|\boldsymbol{\mu}_s\|_{k+1}^2 + \|\pi_s\|_k^2)\mathrm{d}s.$$

Moreover, for $t > 1$, $\tau(t) = 1$ and thus we easily have the same. By arguing similarly, we can bound the second term. Thus, we obtain by Assumption (A) that

$$\|\mathscr{P}_h^s\boldsymbol{\mu} - \boldsymbol{\zeta}_h\|^2 \leq c\left(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2}\right)h^{2k+2}. \tag{3.40}$$

Finally by the inverse estimate (D), we have $h^2\|\nabla(\mathscr{P}_h^s\boldsymbol{\mu} - \boldsymbol{\zeta}_h)\|^2 \leq C(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2})h^{2k+2}$.

For proving (3.35), we will essentially repeat the same steps as we did in proving (3.36). We begin by rewriting $(3.8)_1$ as

$$\begin{aligned}(\partial_t(\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h), \mathbf{v}_h) + \tfrac{1}{Re}(\nabla(\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h), \nabla\mathbf{v}_h) + \tfrac{1}{\epsilon}(\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h, \mathbf{v}_h)_\Gamma \\ = (\partial_t(\mathscr{P}_h^s\mathbf{u} - \mathbf{u}), \mathbf{v}_h) + \tfrac{1}{\epsilon}(\widetilde{\mathbf{e}}_g, \mathbf{v}_h)_\Gamma \quad \mathbf{v}_h \in \mathbf{V}_h.\end{aligned} \tag{3.41}$$

Now setting $\mathbf{v}_h = \sigma(t)(\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h)$ and integrating with respect to time t yields

$$\begin{aligned}\sigma(t)\|\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h\|^2 + \int_0^t \frac{\sigma(t)}{Re}\|\nabla(\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h)\|^2\,\mathrm{d}s + \int_0^t \frac{\sigma(t)}{\epsilon}\|\mathscr{P}_h^s\mathbf{u} - \boldsymbol{\xi}_h\|_{0,\Gamma}^2\,\mathrm{d}s \\ \leq c\int_0^t \frac{\sigma^2}{\partial_s\sigma}\|\partial_t(\mathscr{P}_h^s\mathbf{u} - \mathbf{u})\|^2\,\mathrm{d}s + c\int_0^t \frac{\sigma(t)}{\epsilon^3}\|\widetilde{\mathbf{e}}_{\boldsymbol{\mu}}\|_{0,\Gamma}^2\,\mathrm{d}s \\ + c\int_0^t \mathrm{e}^{\alpha\, s}\|\mathscr{P}_h^s\mathbf{u} - \mathbf{u}\|^2\,\mathrm{d}s + c\int_0^t \mathrm{e}^{\alpha\, s}\,\|\widetilde{\mathbf{e}}_u\|^2\,\mathrm{d}s.\end{aligned} \tag{3.42}$$

Thus by Lemmas 3.2 and 3.3, it follows from (3.42) that

$$\|P_h^s\mathbf{u} - \boldsymbol{\xi}_h\|^2 \leq C\left(1 + \frac{h}{\epsilon} + \frac{h^2}{\epsilon^2} + \frac{h^3}{\epsilon^3}\right)h^{2k+2}. \tag{3.43}$$

The desired result now follows by assuming $\epsilon = \mathcal{O}(h)$. □

Lemma 3.5. *Suppose that Assumptions* (A), (B) *and* (D) *hold and* $\epsilon = \mathcal{O}(h)$. *Then there exists a constant* C, *independent of* h, *such that the error* $\widetilde{\mathbf{e}}_u := \mathbf{u} - \boldsymbol{\xi}_h$ *and* $\widetilde{\mathbf{e}}_\mu := \boldsymbol{\mu} - \boldsymbol{\zeta}_h$ *satisfies*

$$\begin{aligned}\|\mathbf{u} - \boldsymbol{\xi}_h\| + h\|\nabla(\mathbf{u} - \boldsymbol{\xi}_h)\| \leq Ch^{k+1} \\ \|\boldsymbol{\mu} - \boldsymbol{\zeta}_h\| + h\|\nabla(\boldsymbol{\mu} - \boldsymbol{\zeta}_h)\| \leq Ch^{k+1}.\end{aligned} \tag{3.44}$$

Proof. By combining Lemmas 3.3 and 3.4, we obtain the required result. □

We shall next state and prove the main result of this section regarding the error in the semi-discrete velocity and its adjoint.

Theorem 3.6. *Let the Assumptions* (A), (B) *and* (D) *hold true, and let* $\epsilon \in (0, Re)$ *and* $\epsilon = \mathcal{O}(h)$ *be satisfied. Further, let the initial velocity* $\mathbf{u}_{0h}$ *satisfy*

$$\|\mathbf{u}_0 - \mathbf{u}_{0h}\| \leq ch^{k+1}\|\mathbf{u}_0\|_{k+1}.$$

Then there exists a constant C *depending on* T *and independent of* h *such that the following estimate holds:*

$$\|(\mathbf{u}, \boldsymbol{\mu}) - (\mathbf{u}_h, \boldsymbol{\mu}_h)\| + h\|\nabla((\mathbf{u}, \boldsymbol{\mu}) - (\mathbf{u}_h, \boldsymbol{\mu}_h))\| \leq Ch^{k+1}.$$

Proof. First notice that $\widehat{\mathbf{e}}_u$ and $\widehat{\mathbf{e}}_\mu$ satisfy

$$(\partial_t \widehat{\mathbf{e}}_u, \mathbf{v}_h) + \frac{1}{Re}(\nabla \widehat{\mathbf{e}}_u, \nabla \mathbf{v}_h) + \frac{1}{\epsilon}(\widehat{\mathbf{e}}_u - \widehat{\mathbf{e}}_g, \mathbf{v}_h)_\Gamma = c(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) - c(\mathbf{u}, \mathbf{u}, \mathbf{v}_h) =: \aleph_1(\mathbf{v}_h) \tag{3.45}$$

and

$$\begin{aligned} -(\partial_t \widehat{\mathbf{e}}_{\boldsymbol{\mu}}, \mathbf{w}_h) + \frac{1}{Re}(\nabla \widehat{\mathbf{e}}_{\boldsymbol{\mu}}, \nabla \mathbf{w}_h) + \frac{1}{\epsilon}(\widehat{\mathbf{e}}_{\boldsymbol{\mu}}, \mathbf{w}_h)_\Gamma &= (c(\mathbf{u}_h, \mathbf{w}_h, \boldsymbol{\mu}_h) - c(\mathbf{u}, \mathbf{w}_h, \boldsymbol{\mu})) \\ &+ (c(\mathbf{w}_h, \mathbf{u}_h, \boldsymbol{\mu}_h) - c(\mathbf{w}_h, \mathbf{u}, \boldsymbol{\mu})) - \delta(\nabla \mathbf{e}_u, \nabla \mathbf{w}_h) - \delta(\mathbf{e}_u, \mathbf{w}_h) \\ &=: \aleph_2(\mathbf{w}_h) + \aleph_3(\mathbf{w}_h) - \delta(\nabla \mathbf{e}_u, \nabla \mathbf{w}_h) - \delta(\mathbf{e}_u, \mathbf{w}_h). \end{aligned} \tag{3.46}$$

Since

$$\aleph_1(\mathbf{v}_h) = -c(\mathbf{u}, \widetilde{\mathbf{e}}_u, \mathbf{v}_h) - c(\widetilde{\mathbf{e}}_u, \boldsymbol{\xi}_h, \mathbf{v}_h) - c(\widehat{\mathbf{e}}_u, \boldsymbol{\xi}_h, \mathbf{v}_h) - c(\mathbf{u}_h, \widehat{\mathbf{e}}_u, \mathbf{v}_h) =: \sum_{i=1}^{4} \aleph_{1,i}(\mathbf{v}_h),$$

$$\aleph_2(\mathbf{w}_h) = -c(\widehat{\mathbf{e}}_u, \mathbf{w}_h, \boldsymbol{\zeta}_h) - c(\widetilde{\mathbf{e}}_u, \mathbf{w}_h, \boldsymbol{\zeta}_h) - c(\mathbf{u}, \mathbf{w}_h, \widetilde{\mathbf{e}}_\mu) - c(\mathbf{u}_h, \mathbf{w}_h, \widehat{\mathbf{e}}_\mu) =: \sum_{i=1}^{4} \aleph_{2,i}(\mathbf{w}_h),$$

and

$$\aleph_3(\mathbf{w}_h) = 2c(\mathbf{w}_h, \mathbf{u}, \widehat{\mathbf{e}}_\mu) + c(\mathbf{w}_h, \mathbf{u}, \widetilde{\mathbf{e}}_\mu) - c(\mathbf{w}_h, \widehat{\mathbf{e}}_u, \boldsymbol{\zeta}_h) - c(\mathbf{w}_h, \widetilde{\mathbf{e}}_u, \boldsymbol{\zeta}_h) + c(\mathbf{w}_h, \mathbf{u}_h, \widehat{\mathbf{e}}_\mu) =: \sum_{i=1}^{5} \aleph_{3,i}(\mathbf{w}_h)$$

setting $(\mathbf{v}_h, \mathbf{w}_h) = (\widehat{\mathbf{e}}_u, \widehat{\mathbf{e}}_\mu)$ in (3.45)–(3.56) yields

$$\begin{cases} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widehat{\mathbf{e}}_u\|^2 + \frac{1}{Re}\|\nabla \widehat{\mathbf{e}}_u\|^2 + \frac{1}{\epsilon}\|\widehat{\mathbf{e}}_u\|_{0,\Gamma}^2 \leq \aleph_1(\widehat{\mathbf{e}}_u) + \frac{1}{\epsilon^2}\|\widehat{\mathbf{e}}_u\|_1 \|\widehat{\mathbf{e}}_\mu\|_{-\frac{1}{2},\Gamma} \\ -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widehat{\mathbf{e}}_\mu\|^2 + \frac{1}{Re}\|\nabla \widehat{\mathbf{e}}_\mu\|^2 + \frac{1}{\epsilon}\|\widehat{\mathbf{e}}_\mu\|_{0,\Gamma}^2 = \aleph_2(\widehat{\mathbf{e}}_\mu) + \aleph_3(\widehat{\mathbf{e}}_\mu) + \delta(\nabla \mathbf{e}_u, \nabla \widehat{\mathbf{e}}_\mu) \\ \qquad + \delta(\mathbf{e}_u, \widehat{\mathbf{e}}_\mu), \end{cases} \tag{3.47}$$

where we have used the relation $\widehat{\mathbf{e}}_g = -\frac{1}{\beta\epsilon}\widehat{\mathbf{e}}_\mu$. Using Young's inequality, we obtain

$$-\delta\|\nabla \widehat{\mathbf{e}}_u\|\|\nabla \widehat{\mathbf{e}}_\mu\| \geq -\frac{\delta}{2\eta_1}\|\widehat{\mathbf{e}}_u\|_1^2 - \eta_1 \frac{\delta}{2}\|\widehat{\mathbf{e}}_\mu\|_1^2$$

for some $\eta_1 \geq \frac{\delta Re}{2\widehat{C}}$, where $\widehat{C}$ is as defined in (2.1). Employing this inequality in $(3.47)_2$ and using (2.1), we obtain

$$\begin{aligned} -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widehat{\mathbf{e}}_\mu\|^2 + \widehat{\eta}_2\|\widehat{\mathbf{e}}_\mu\|_1^2 \leq & \aleph_2(\widehat{\mathbf{e}}_\mu) + \aleph_3(\widehat{\mathbf{e}}_\mu) + \frac{\delta\eta_1}{2}\|\widehat{\mathbf{e}}_u\|_1^2 + \frac{\widehat{\eta}_2}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2 + c\|\nabla \widetilde{\mathbf{e}}_u\|^2 \\ & + \frac{\delta}{2}\|(\widetilde{\mathbf{e}}_u, \widetilde{\mathbf{e}}_\mu)\|^2 + \delta\|\widehat{\mathbf{e}}_\mu\|^2. \end{aligned} \tag{3.48}$$

where $\widehat{\eta}_2 := \frac{\widehat{C}}{Re} - \frac{\delta}{2\eta_1}$. We use inverse inequality and Young's inequality to estimate the second term on the right-hand side of $(3.47)_1$ as follows

$$\begin{aligned} \frac{1}{\epsilon^2}\|\widehat{\mathbf{e}}_u\|_1 \|\widehat{\mathbf{e}}_\mu\|_{-\frac{1}{2},\Gamma} &\leq \frac{Ch}{\epsilon^2}\|\widehat{\mathbf{e}}_u\|_1 \|\widehat{\mathbf{e}}_\mu\|_1 \\ &\leq \frac{Ch^2}{\epsilon^4}\|\widehat{\mathbf{e}}_u\|_1^2 + \frac{\widehat{\eta}_2}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2 \leq \frac{Ch^2}{\epsilon^8}\|\widehat{\mathbf{e}}_u\|^2 + \frac{\delta\eta_1}{2\eta_2}\|\widehat{\mathbf{e}}_u\|_1^2 + \frac{\widehat{\eta}_2}{8\eta_2}\|\widehat{\mathbf{e}}_\mu\|_1^2. \end{aligned} \tag{3.49}$$

The nonlinear terms $\aleph_1 - \aleph_3$ and other terms in the right-hand side of (3.47) can be estimated as below. First by skew-symmetry $\aleph_{1,4}(\widehat{\mathbf{e}}_u) = \aleph_{2,4}(\widehat{\mathbf{e}}_\mu) = 0$. Estimating $\aleph_{1,1}(\widehat{\mathbf{e}}_u), \aleph_{1,2}(\widehat{\mathbf{e}}_u)$ and $\aleph_{1,3}(\widehat{\mathbf{e}}_u)$ using (A.1), (A.2) and (A.5), respectively, and collecting the estimates, we obtain

$$\aleph_1(\widehat{\mathbf{e}}_u) \leq \frac{\delta\eta_1}{8\eta_2}\|\widehat{\mathbf{e}}_u\|_1^2 + c\|\widetilde{\mathbf{e}}_u\|_1^2[\|\mathbf{u}\|_2^2 + \|\mathbf{u}\|_1^2 + \|\boldsymbol{\xi}_h\|_\infty^2 + \|\nabla\boldsymbol{\xi}_h\|^2] + \|\widehat{\mathbf{e}}_u\|^2[\|\boldsymbol{\xi}_h\|_\infty^2 + \|\nabla\boldsymbol{\xi}_h\|^2]. \tag{3.50}$$

Estimating $\aleph_{2,1}(\widehat{\mathbf{e}}_\mu)$, $\aleph_{2,2}(\widehat{\mathbf{e}}_\mu)$ and $\aleph_{2,3}(\widehat{\mathbf{e}}_\mu)$ using (A.2), (A.7) and (A.4), respectively, and collecting the estimates, we obtain

$$\begin{aligned}\aleph_2(\widehat{\mathbf{e}}_\mu) \leq{}& \tfrac{\widehat{\eta_2}}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2 + \tfrac{\delta\eta_1}{8}\|\widehat{\mathbf{e}}_u\|_1^2 + \|\widetilde{\mathbf{e}}_u\|_1^2[\|\boldsymbol{\zeta}_h\|_\infty^2 + \|\nabla\boldsymbol{\zeta}_h\|^2 + c]\\ &+ c\|\widehat{\mathbf{e}}_u\|^2[\|\boldsymbol{\zeta}_h\|_\infty^2 + \|\nabla\boldsymbol{\zeta}_h\|^2] + c\|\mathbf{u}\|_2^2[\|\widetilde{\mathbf{e}}_\mu\|^2 + \|\widehat{\mathbf{e}}_\mu\|^2] + c\|\widehat{\mathbf{e}}_\mu\|^2.\end{aligned} \tag{3.51}$$

Similarly, estimating $\aleph_{3,1}(\widehat{\mathbf{e}}_\mu)$, $\aleph_{3,2}(\widehat{\mathbf{e}}_\mu)$, $\aleph_{3,3}(\widehat{\mathbf{e}}_\mu)$, $\aleph_{3,4}(\widehat{\mathbf{e}}_\mu)$ and $\aleph_{3,5}(\widehat{\mathbf{e}}_\mu)$ using (A.5), (A.6), (A.3), (A.3) and (A.8), respectively, and collecting the estimates, we obtain

$$\begin{aligned}\aleph_3(\widehat{\mathbf{e}}_\mu) \leq{}& \tfrac{\widehat{\eta_2}}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2 + \|\widehat{\mathbf{e}}_\mu\|^2[\|\mathbf{u}_h\|_{L^4}^4 + \|\nabla\mathbf{u}_h\|^2 + \|\nabla\mathbf{u}\|_{L^4}^2 + \|\mathbf{u}\|_2^2 + \|\boldsymbol{\zeta}_h\|_\infty^2 + \|\nabla\boldsymbol{\zeta}_h\|^2]\\ &+ \tfrac{\delta\eta_1}{8}\|\widehat{\mathbf{e}}_u\|_1^2 + \|\widetilde{\mathbf{e}}_\mu\|^2[\|\nabla\mathbf{u}\|_{L^4}^2 + \|\mathbf{u}\|_2^2] + C[\|\widetilde{\mathbf{e}}_u\|_1^2 + \|\widehat{\mathbf{e}}_u\|^2].\end{aligned} \tag{3.52}$$

Multiplying $(3.47)_1$ by η_1 and inserting (3.50)$-$(3.52) in the resulting equation and also in (3.48), we obtain, respectively,

$$\begin{cases}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\eta_2\|\widehat{\mathbf{e}}_u\|^2) + \frac{\eta_2\widehat{C}}{Re}\|\widehat{\mathbf{e}}_u\|_1^2 \leq \frac{\delta\eta_1}{4}\|\widehat{\mathbf{e}}_u\|_1^2 + H_1\|\widetilde{\mathbf{e}}_u\|_1^2 + (\frac{Ch^2}{\epsilon^8} + H_2)\|\widehat{\mathbf{e}}_u\|^2\\ \qquad\qquad + \frac{\widehat{\eta_2}}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2,\\ -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widehat{\mathbf{e}}_\mu\|^2 + \widehat{\eta}_2\|\widehat{\mathbf{e}}_\mu\|_1^2 \leq \frac{3\widehat{\eta_2}}{8}\|\widehat{\mathbf{e}}_\mu\|_1^2 + \frac{3\delta\eta_1}{4}\|\widehat{\mathbf{e}}_u\|_1^2 + H_3\|(\widetilde{\mathbf{e}}_u, \widetilde{\mathbf{e}}_\mu)\|_1^2\\ \qquad\qquad + H_4\|(\widehat{\mathbf{e}}_u, \widehat{\mathbf{e}}_\mu)\|^2 + H_5\|\widetilde{\mathbf{e}}_\mu\|^2 + \frac{\delta}{2}\|\widetilde{\mathbf{e}}_u\|^2,\end{cases} \tag{3.53}$$

where

$$\begin{aligned}&H_1(t) := C[\|\mathbf{u}\|_2^2 + \|\mathbf{u}\|_1^2 + \|\boldsymbol{\xi}_h\|_\infty^2 + \|\nabla\boldsymbol{\xi}_h\|^2],\\ &H_2(t) := C[\|\boldsymbol{\xi}_h\|_\infty^2 + \|\nabla\boldsymbol{\xi}_h\|^2],\\ &H_3(t) := C[1 + \|\boldsymbol{\zeta}_h\|_\infty^2 + \|\nabla\boldsymbol{\zeta}_h\|^2],\\ &H_4(t) := C[\|\boldsymbol{\zeta}_h\|_\infty^2 + \|\nabla\boldsymbol{\zeta}_h\|^4 + \|\mathbf{u}_h\|_{L^4}^4 + \|\nabla\mathbf{u}_h\|^2 + \|\nabla\mathbf{u}\|_{L^4}^2 + \|\mathbf{u}\|_2^2 + 1],\\ &H_5(t) := C[\|\nabla\mathbf{u}\|_{L^4}^2 + \|\mathbf{u}\|_2^2]\end{aligned}$$

and $\eta_2 > \frac{Re\delta\eta_1}{\widehat{C}}$. Notice by virtue of the regularity properties of the solutions and Lemma 3.5 that $\|\nabla(\boldsymbol{\zeta}_h, \boldsymbol{\xi}_h)\| < C$. Moreover, it can be shown using inverse estimate, approximation properties and (3.44) that

$$\|\boldsymbol{\xi}\|_\infty \leq C\|\mathbf{u}\|_2 + K_1h^{1/2} \qquad \|\boldsymbol{\zeta}\|_\infty \leq C\|\mu\|_2 + K_2h^{1/2},$$

see for *e.g.* [14]. Therefore we easily see by Assumption (A) and Lemma 3.1 that $H_i(t) \in L^1(0,T), i = 1,2,3,4,5$. Therefore applying Gronwall's inequality, Lemmas 3.2 and 3.5, and setting $\epsilon = \mathcal{O}(h^{\frac{1}{4}})$, it follows that

$$\eta_1\|\widehat{\mathbf{e}}_u\|^2 + \|\widehat{\mathbf{e}}_\mu\|^2 + \widehat{\eta_1}\int_0^T \|\widehat{\mathbf{e}}_u\|_1^2\,\mathrm{d}s + \widehat{\eta_2}\int_0^T \|\widehat{\mathbf{e}}_\mu\|_1^2\,\mathrm{d}s \leq Ch^{2k+2},$$

where $\widehat{\eta_1} := \eta_1 - \frac{\delta Re}{2\widehat{C}}$ and $\widehat{\eta_2} := \eta_2 - \frac{Re\delta\eta_1}{\widehat{C}}$. Hence the desired result follows from the triangle inequality and inverse inequality. □

Let us next obtain some estimates in preparation for proving error estimates for the pressure and adjoint pressure approximations.

Lemma 3.7. *Suppose that Assumptions* (B.1)–(B.3) *hold and* $\epsilon = \mathcal{O}(h)$. *Then the time derivative of the errors* $\mathbf{e}_u$ *and* $\mathbf{e}_\mu$ *satisfy*

$$\|\partial_t \mathbf{e}_u\| \le K_1 \tau(t)^{-\frac{1}{2}} h^k \quad \text{and} \quad \|\partial_t \mathbf{e}_\mu\| \le K_2 \tau(t)^{-\frac{1}{2}} h^k,$$

$t > 0$, *for some constants* K_1 *and* K_2, *where* $\tau(t) := \min\{1, t\}$.

Proof. Let us first prove that

$$\int_0^t \mathrm{e}^{\alpha\, s} \|\partial_t \mathbf{e}_u\|^2 \,\mathrm{d}s \le C_1 h^{2k} \quad \text{and} \quad \int_0^t \mathrm{e}^{\alpha\, s} \|\partial_t \mathbf{e}_\mu\|^2 \,\mathrm{d}s \le C_2 h^{2k}. \tag{3.54}$$

To this end, we split the integrals as follows

$$\int_0^t \mathrm{e}^{\alpha\, s} \|\partial_t \mathbf{e}_u\|^2 \,\mathrm{d}s = \int_0^{h^{k+1}} \mathrm{e}^{\alpha\, s} \, \|\partial_t \mathbf{e}_u\|^2 \,\mathrm{d}s + \int_{h^{k+1}}^t \mathrm{e}^{\alpha\, s} \, \|\partial_t \mathbf{e}_u\|^2 \,\mathrm{d}s = I_1^u + I_2^u$$

and

$$\int_0^t \mathrm{e}^{\alpha\, s} \|\partial_t \mathbf{e}_\mu\|^2 \,\mathrm{d}s = \int_0^{h^{k+1}} \mathrm{e}^{\alpha\, s} \, \|\partial_t \mathbf{e}_\mu\|^2 \,\mathrm{d}s + \int_{h^{k+1}}^t \mathrm{e}^{\alpha\, s} \, \|\partial_t \mathbf{e}_\mu\|^2 \,\mathrm{d}s = I_1^\mu + I_2^\mu.$$

Note that since

$$I_1^u \le c\, h^{k+1} \mathrm{e}^{\alpha h^{k+1}} (\|\partial_t \mathbf{u}\|^2 + \|\partial_t \mathbf{u}_h\|^2) \le K\, h^{k+1}$$

and similarly $I_1^\mu \le c\, h^{k+1}$ it is enough to estimate I_2^u and I_2^μ. Notice that $\mathbf{e}_u$ and $\mathbf{e}_\mu$ satisfy

$$\begin{cases} (\partial_t \mathbf{e}_u, \mathbf{v}_h) + \frac{1}{Re}(\nabla \mathbf{e}_u, \nabla \mathbf{v}_h) + b(\mathbf{v}_h, e_p) + \frac{1}{\epsilon}(\mathbf{e}_u, \mathbf{v}_h)_\Gamma = \frac{1}{\epsilon}(\mathbf{e}_g, \mathbf{v}_h)_\Gamma + \widehat{\Lambda}_h^u(\mathbf{v}_h), \\ -(\partial_t \mathbf{e}_\mu, \mathbf{w}_h) + \frac{1}{Re}(\nabla \mathbf{e}_\mu, \nabla \mathbf{w}_h) + b(\mathbf{w}_h, e_\pi) + \frac{1}{\epsilon}(\mathbf{e}_\mu, \mathbf{w}_h)_\Gamma = \delta(\nabla \mathbf{e}_u, \nabla \mathbf{w}_h) \\ \qquad\qquad + \delta(\mathbf{e}_u, \mathbf{w}_h) + \widehat{\Lambda}_h^\mu(\mathbf{w}_h), \end{cases} \tag{3.55}$$

where $\widehat{\Lambda}_h^u(\mathbf{v}_h) := -c(\mathbf{e}_u, \mathbf{u}, \mathbf{v}_h) - c(\mathbf{u}, \mathbf{e}_u, \mathbf{v}_h) + c(\mathbf{e}_u, \mathbf{e}_u, \mathbf{v}_h)$ and $\widehat{\Lambda}_h^\mu(\mathbf{w}_h) := c(\mathbf{u}, \mathbf{w}_h, \mathbf{e}_\mu) - c(\mathbf{e}_u, \mathbf{w}_h, \mathbf{e}_\mu) + c(\mathbf{e}_u, \mathbf{w}_h, \mu) + c(\mathbf{w}_h, \mathbf{u}, \mathbf{e}_\mu) - c(\mathbf{w}_h, \mathbf{e}_u, \mathbf{e}_\mu) + c(\mathbf{w}_h, \mathbf{e}_u, \mu)$. Setting $\mathbf{v}_h = \partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)$ and $\mathbf{w}_h = \partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)$ in $(3.55)_1$ and $(3.55)_2$, respectively, we obtain

$$\begin{aligned} 2\|\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)\|^2 &+ \frac{1}{Re}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)\|^2 + \frac{1}{\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h\|_{0,\Gamma}^2 \\ &= \frac{2}{\epsilon}(\mathbf{e}_g, \partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)_\Gamma + 2\widehat{\Lambda}_h^u(\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)) \\ &\quad + 2(\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}), \partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)), \\ -2\|\partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)\|^2 &+ \frac{1}{Re}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)\| + \frac{1}{\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h\|_{0,\Gamma}^2 \\ &= 2\delta(\nabla \mathbf{e}_u, \nabla\partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)) \\ &\quad + 2\delta(\mathbf{e}_u, \partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)) + 2\widehat{\Lambda}_h^\mu(\partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)) \\ &\quad - 2(\partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}), \partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h)). \end{aligned}$$

We next multiply these equations by $\mathrm{e}^{\alpha\, t}$, integrate with respect to time from h^2 to t and estimate the terms as usual. For example, we estimate $\widehat{\Lambda}_h^u(\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h))$ and $\widehat{\Lambda}_h^u(\partial_t(\mathscr{P}_h^s \boldsymbol{\mu} - \boldsymbol{\mu}_h))$ using Hölder's and Sobolev inequality as follows:

$$|\widehat{\Lambda}_h^u| \le C[\|\mathbf{e}_u\|\|\mathbf{u}\|_2\|\nabla\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)\| + \|\mathbf{u}\|_2\|\mathbf{e}_u\|_1\|\|\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)\| + \|\nabla \mathbf{e}_u\|^2\|\nabla\partial_t(\mathscr{P}_h^s \mathbf{u} - \mathbf{u}_h)\|] \tag{3.56}$$

and

$$\begin{aligned}|\widehat{\Lambda}_h^\mu| \le C[&\|\mathbf{u}\|_2\|\mathbf{e}_\mu\|\|\nabla\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\| + \|\mathbf{u}\|_2\|\nabla\mathbf{e}_\mu\|\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\| \\ &+\|\nabla\mathbf{e}_\mu\|\|\nabla\mathbf{e}_u\|\|\nabla\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\| + \|\boldsymbol{\mu}\|_2\|\mathbf{e}_u\|\|\nabla\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\| \\ &+\|\boldsymbol{\mu}\|_2\|\nabla\mathbf{e}_u\|\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\|].\end{aligned} \tag{3.57}$$

Therefore by applying the inverse inequality to (3.56) and (3.57) and using Theorem 3.6 yields

$$\begin{aligned}|\widehat{\Lambda}_h^u| &\le C[\|\mathbf{u}\|_2\|\partial_t(\mathscr{P}_h^s\mathbf{u}-\mathbf{u}_h)\|(h^{-1}\|\mathbf{e}_u\| + \|\nabla\mathbf{e}_u\|) + \|\nabla\mathbf{e}_u\|^2\|\nabla\partial_t(\mathscr{P}_h^s\mathbf{u}-\mathbf{u}_h)\|] \\ &\le C\|\nabla\mathbf{e}_u\|^2 + \tfrac14\|\partial_t(\mathscr{P}_h^s\mathbf{u}-\mathbf{u}_h)\|^2\end{aligned} \tag{3.58}$$

and

$$\begin{aligned}|\widehat{\Lambda}_h^\mu| \le& C[\|\mathbf{u}\|_2\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\|(h^{-1}\|\mathbf{e}_\mu\| + \|\nabla\mathbf{e}_\mu\|) + \|\boldsymbol{\mu}\|_2\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\|(h^{-1}\|\mathbf{e}_u\| + \|\nabla\mathbf{e}_u\|) \\ &+ h^{-1}\|\nabla\mathbf{e}_u\|\|\nabla\mathbf{e}_\mu\|\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\|] \\ \le& C[\|\nabla\mathbf{e}_u\|^2 + \|\nabla\mathbf{e}_\mu\|^2] + \frac14\|\partial_t(\mathscr{P}_h^s\boldsymbol{\mu}-\boldsymbol{\mu}_h)\|^2.\end{aligned} \tag{3.59}$$

Employing (3.58) and (3.59) in (3.55), estimating the other terms there as usual, integrating with respect to time and using Lemma 3.3 and Theorem 3.6, we obtain (3.54). The required result can now be obtained by differentiating (3.55) with respect to time, setting $(\mathbf{v}_h, \mathbf{w}_h) = (\sigma(t)\mathscr{P}_h\partial_t\mathbf{e}_u, \sigma(t)\mathscr{P}_h\partial_t\mathbf{e}_\mu)$ and arguing as we did above. □

Theorem 3.8. *Suppose that Assumptions* (A)–(D) *hold, and* $\epsilon \in (0, Re)$ *and* $\epsilon = \mathcal{O}(h)$. *Then there exists constants* K_1 *and* K_2 *such that the following error estimates hold*

$$\|p - p_h\| \le K_1\tau(t)^{-\frac12}h^k \quad \textit{and} \quad \|\pi - \pi_h\| \le K_2\tau(t)^{-\frac12}h^k,$$

$t > 0$, *where* $\tau(t) := \min\{1, t\}$.

Proof. First notice that the error $p - p_h$ satisfy

$$\begin{aligned}b(\mathbf{v}_h, p - p_h) &= -(\partial_t\mathbf{e}_u, \mathbf{v}_h) - \tfrac{1}{Re}(\nabla\mathbf{e}_u, \nabla\mathbf{v}_h) - c(\mathbf{e}_u, \mathbf{u}_h, \mathbf{v}_h) - c(\mathbf{u}, \mathbf{e}_u, \mathbf{v}_h) \\ &\quad -\tfrac{1}{\epsilon}(\mathbf{e}_u, \mathbf{v}_h)_\Gamma + \tfrac{1}{\epsilon}(\mathbf{e}_g, \mathbf{v}_h)_\Gamma \\ &\le C[\|\partial_t\mathbf{e}_u\| + \|\nabla\mathbf{e}_u\| + \tfrac1\epsilon\|\mathbf{e}_u\|_{-\frac12,\Gamma} + \tfrac1\epsilon\|\mathbf{e}_g\|_{-\frac12,\Gamma}]\|\mathbf{v}_h\|_1.\end{aligned} \tag{3.60}$$

Moreover, since

$$\begin{aligned}b(\mathbf{w}_h, \pi - \pi_h) = &-(\partial_t\mathbf{e}_\mu, \mathbf{w}) - \frac{1}{Re}(\nabla\mathbf{e}_\mu, \nabla\mathbf{w}_h) - c(\mathbf{u}, \mathbf{w}_h, \mathbf{e}_\mu) - c(\mathbf{e}_u, \mathbf{w}_h, \mu_h) \\ &- \frac1\epsilon(\mathbf{e}_\mu, \mathbf{w}_h)_\Gamma + \delta(\nabla\mathbf{e}_\mu, \nabla\mathbf{w}_h) + \delta(\mathbf{e}_\mu, \mathbf{w}_h) \\ &- c(\mathbf{w}_h, \mathbf{u}, \mathbf{e}_\mu) - c(\mathbf{w}_h, \mathbf{e}_u, \boldsymbol{\mu}_h),\end{aligned}$$

we have that

$$b(\mathbf{w}_h, \pi - \pi_h) \le C[\|\partial_t\mathbf{e}_\mu\| + \|\nabla\mathbf{e}_\mu\| + \|\nabla\mathbf{e}_u\| + \frac1\epsilon\|\mathbf{e}_\mu\|_{-\frac12,\Gamma}]\|\mathbf{w}_h\|_1. \tag{3.61}$$

It follows from the inf-sup condition and (3.60) that

$$\begin{aligned}\|p - p_h\| &\le C\|p - q_h\| + \sup_{\mathbf{v}_h\in\mathbf{X}_h}\frac{b(p - p_h, \mathbf{v}_h)}{\|\mathbf{v}_h\|_1} \\ &\le C[\|p - q_h\| + \|\partial_t\mathbf{e}_u\| + \|\nabla\mathbf{e}_u\| + \frac1\epsilon\|\mathbf{e}_u\|_{-\frac12,\Gamma} + \frac1\epsilon\|\mathbf{e}_g\|_{-\frac12,\Gamma}],\end{aligned} \tag{3.62}$$

for $q_h \in Q_h$. Combining (3.62), Theorem 3.6 and Lemma 3.7, we obtain $\|p - p_h\| \le K_1\tau(t)^{-\frac12}h^k$. A similar argument with the help of (3.61) leads to $\|\pi - \pi_h\| \le K_2\tau(t)^{-\frac12}h^k$. □

4. Computational results

In this section, we show feasibility and applicability of the penalty method by using it to solve an optimal boundary control problem in wall bounded channel flow. The flow configuration is the backward facing step channel with fixed channel width ratio and Reynolds number. Beyond certain Reynolds number, flow separates near the enlargement due to pressure increase. Subsequently, the flow re-attaches on the bottom wall, and recirculation forms near the corner region. After re-attachment, the flow field fully recovers toward a fully developed Poiseuille flow [1, 9, 20]. It is of interest to alleviate flow separation and wake spread, and thus improving the performance of the fluid system. Below we formulate it as an optimal boundary control problem in which the control that minimizes the cost functional is found using a variable step gradient algorithm, where gradient of the objective function is obtained by solving adjoint equations. The control is effected via small suction and blowing through a slot on the channel boundary. The choice of cost functional or objective functional to meet the control objective of reducing the flow separation and recirculation is not trivial. Here we will consider cost functional as defined in (P) with control $\mathbf{g} = (g, 0)$ and three different choices for Θ, namely, $\Theta_1 := \frac{\delta}{2}[\|\mathbf{u}\|^2 + \|\nabla\mathbf{u}\|^2]$ which corresponds to minimizing kinetic energy in H^1-norm, $\Theta_2(\mathbf{u}) = \frac{\delta}{2}[\|\nabla\times\mathbf{u}\|^2]$ which corresponds to minimization of enstrophy levels in the flow and $\Theta_3 := \frac{\delta}{2}[\|\nabla\mathbf{u} + (\nabla\mathbf{u})^T\|^2]$ which corresponds to minimizing viscous dissipation function.

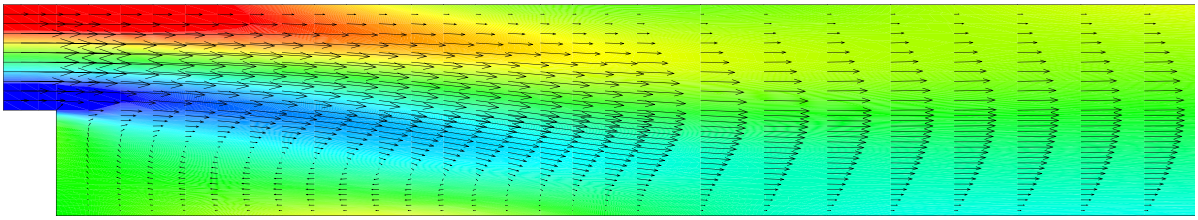
In Figure 1, the downstream channel was defined to have unit height L with a step height and inlet height L/2. The downstream channel length was taken as x=12L. The only non-dimensional parameter of interest, the Reynolds' number is defined by $Re = u_{ave}L/\nu$. At the inflow channel boundary a parabolic velocity profile is prescribed, *i.e.* $u(x=0,\ 1/2 \le y \le 1) = 24(y-1/2)(1-y)$, $v(x=0,\ 1/2 \le y \le 1) = 0$, which produces a maximum inflow velocity of $u_{max} = 3/2$ and an average velocity of $u_{ave} = 1$. On the solid walls the no-slip condition ($\mathbf{u} = 0$) is imposed. At the outflow, the pseudo stress-free condition,

$$-p + \frac{1}{Re}\frac{\partial u}{\partial x} = 0 \quad \text{and} \quad \frac{\partial v}{\partial x} = 0,$$

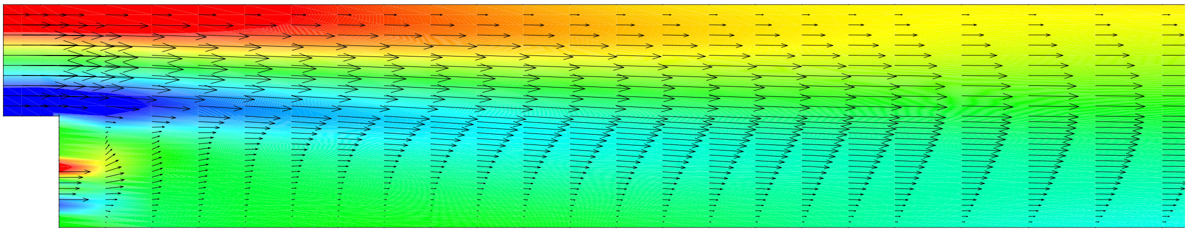
is applied [9].

The computational grid was non-uniform in both the stream-wise and cross-flow coordinate directions. The velocity and adjoint velocity were approximated by piecewise quadratic polynomials while pressure and adjoint pressure were approximated by piecewise linear polynomials. All the variables were defined on the same triangulation and on each triangle the degrees of freedom for quadratic elements were the function values at the vertices and midpoints of each edge; the degrees of freedom for linear elements were the function values at the vertices. This particular choice of finite element spaces for velocity and pressure satisfies the discrete inf-sup condition. The time discretization is carried out using a second-order extrapolated backward difference formula (BDF2) [22]. A fine grid was used in regions where sharp variations in velocities were expected. All the computations were done with 45×45 grid and a time step size $\Delta t = 1/200$ for the Reynolds' number 200. The flow separates at the corner of the step and a recirculation forms. After the re-attachment of the lower wall eddy, the flow slowly recovers towards a fully developed Poiseuille flow. The predicted re-attachment point on the lower wall was five step-heights downstream. The resulting steady flow field is given in Figure 1.

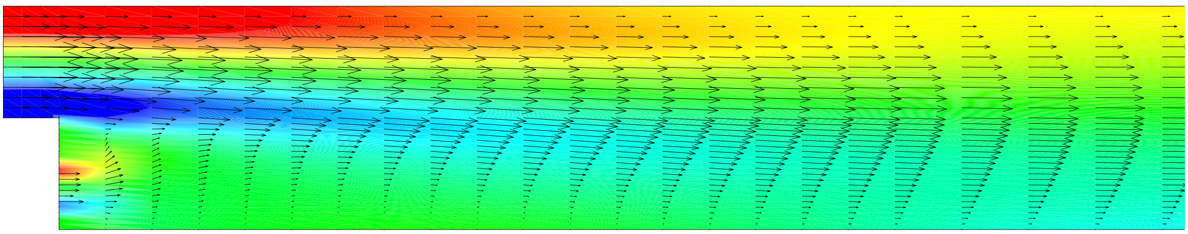
The selection of the portion of the boundary Γ_c, where control is applied, is crucial for the control effectiveness. To this end, first the vertical part of the step was selected for actuator placement. This choice was motivated by the fact that if one wants maximum influence in the flow, then the control has to be applied in the vicinity of the point of separation and wake region. To find where exactly on this part of the channel boundary the actuator should be placed, this portion of the step was divided into four parts of equal size and named them as Slot 1, Slot 2, Slot 3 and Slot 4, where Slot i is defined as $(i-1)/8 \le y \le i/8$, $i = 1, 2, 3, 4$. In order to determine the best position for actuation, the control for each of these cases was computed and results were compared.



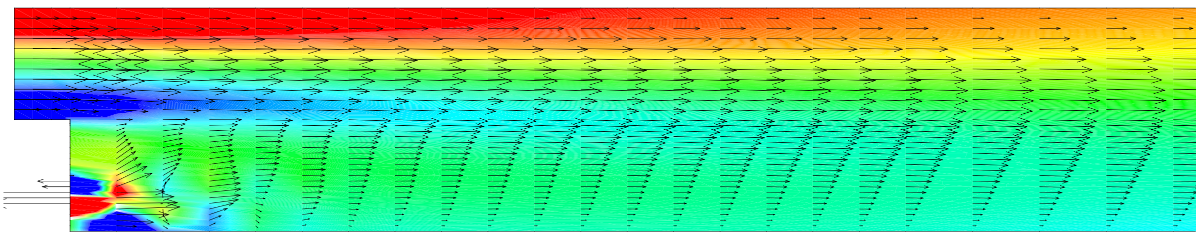
(a)



(b)



(c)



(d)

FIGURE 1. Uncontrolled and controlled velocity fields at $t = 10$ with embedded vorticity contours. (a) Baseline velocity field, (b) Controlled velocity field with actuation at slot 2 and with cost function Θ_1. (c) Controlled velocity field with actuation at slot 2 and with cost function Θ_2. Controlled velocity field with actuation at slot 2 and with cost function Θ_3.

The control that minimizes the cost functional was found using a variable step gradient algorithm. Each iteration of the gradient algorithm requires sequential solution of the state equation (2.6) and adjoint equation (2.8). Adjoint equations were discretized using the same space-time discretization scheme as the one used for the state equations. As these two can not be solved simultaneously in practice, the state equations are solved marching

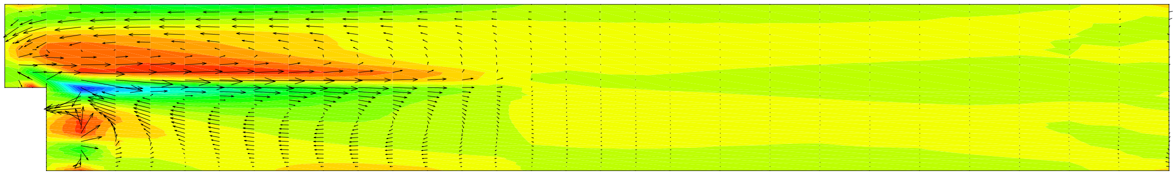


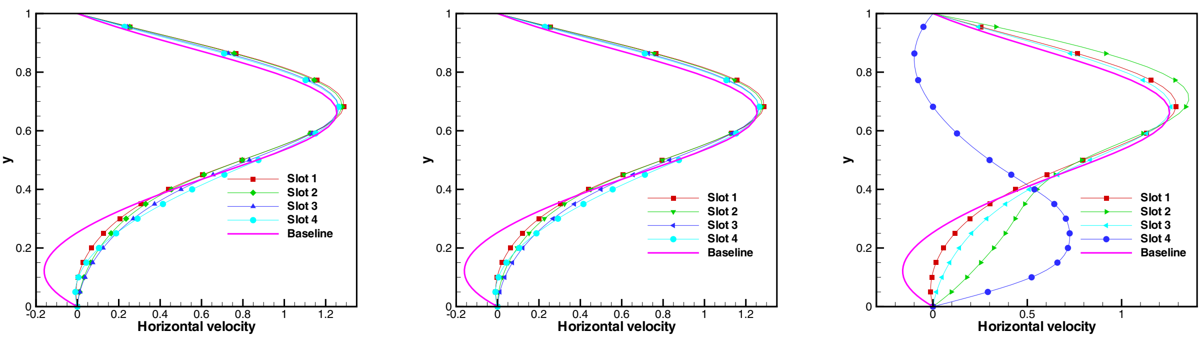
FIGURE 2. Adjoint velocity fields at $t = 10$ with embedded adjoint vorticity contours. With actuation at slot 2 and with cost function Θ_1.

TABLE 1. The $L^2(\Omega)$-norm of vorticity of optimal solution.

ϵ_i	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$\|\nabla \times \mathbf{u}_h^\epsilon\|^2$	0.62	0.58	0.54	0.53	0.52	0.52	0.52

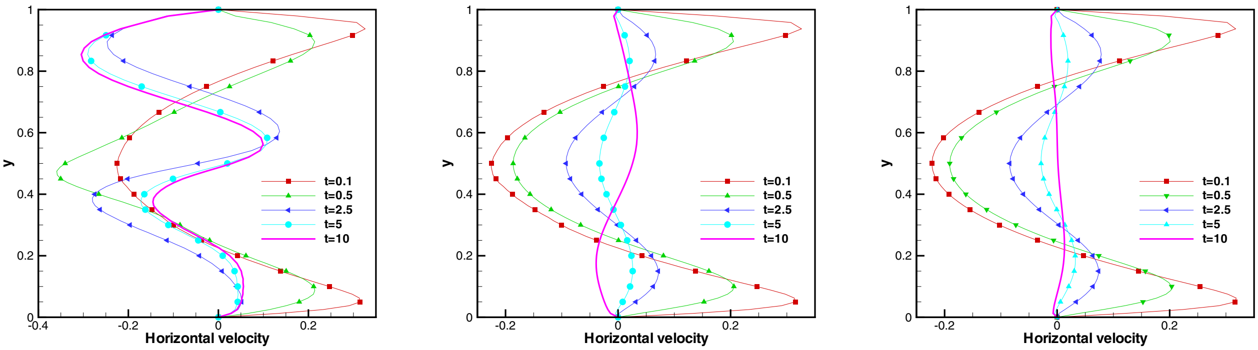
forward in time starting from the initial conditions and adjoint equations are solved marching backward in time from the final conditions at $t = 10$.

Figure 1a shows the baseline velocity field at time $t = 10$. Figure 2b, 2c and 2d show controlled velocity fields computed with cost functions corresponding to Θ_1, Θ_2 and Θ_3, respectively. Figures 1b–1d illustrates the suppression of flow reversal present in the uncontrolled flow shown in Figure 1a. As indicated in the flow fields, separation has been effectively eliminated by the optimal blowing control. Substantial reduction in the recirculation bubble is also seen. The re-attachment length has been reduced by more than 99% compared to the uncontrolled case. Moreover, the results are almost independent of the choice of cost functional employed. In order to verify the convergence of solutions of penalized optimal control problem to that of the Dirichlet control problem with respect to ϵ, numerical experiments were carried out in which optimal solutions were computed for a sequence of ϵ values. In Table 1, $L^2(\Omega)$-norm of vorticity of optimal solution is shown for a sequence of ϵ values. As can be seen in Table 1, convergence do occur as $\epsilon \to 0$. Figure 2 shows the adjoint velocity field associated with vorticity cost function at time $t = 10$ when the actuation is on Slot 2. The adjoint velocity field seems to settles down to steady state when integrated backwards in time and it concentrates entirely in the wake region. As shown in Figure 2, the main feature of adjoint field is a pair of elongated vortices upstream of the channel. The adjoint velocity field propagates upstream while being diffused by the high viscosity. The upstream propagation of adjoint field is due to backward in time integration of the adjoint equations. These observations are also consistent with adjoint fields computed with the other two cost functions. Figure 3 shows a comparison



(a) (b) (c)

FIGURE 3. A comparison of cross-channel profiles of horizontal velocity component of the base line flow with that of the controlled flow corresponding to each actuator position. (a) With cost function Θ_1, (b) With cost function Θ_2, (c) With cost function Θ_3.



(a) (b) (c)

FIGURE 4. Cross-channel profiles of horizontal adjoint velocity component for five different time instances at three different spatial locations. (a) Station 1: $x = 1$, (b) Station 2: $x = 4$ and (c) Station 3: $x = 8$.

of cross channel profiles of horizontal velocity component of the baseline flow with that of the controlled flow corresponding to each actuator position and cost function. These results and others not reported here clearly indicated that Slot 2 is the best place for actuator and the second best place is Slot 1. Moreover, the optimal actuator location is independent of the cost functional employed in the control problem. Figure 4 shows the cross-channel profiles of horizontal adjoint velocity component for five different time instances at three different spatial locations. As can be seen the adjoint field has a larger magnitude at smaller t. The magnitude of the adjoint increases rapidly as t decreases.

APPENDIX A

In this section, we list some of the estimates regarding the trilinear forms that appear in the governing equations and related equations. These results are used in energy arguments throughout the paper.

Lemma A.1.

(i) *For any* $\epsilon > 0$ *and* $\mathbf{u} \in \mathbf{H}^2(\Omega)$, $\mathbf{v}, \mathbf{w} \in H^1(\Omega)$, *there exists a positive constant* C_ϵ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\mathbf{w}\|_1^2 + C_\epsilon \|\mathbf{u}\|_2^2 [\|\mathbf{v}\| + \|\nabla \mathbf{v}\|^2]. \tag{A.1}$$

(ii) *For any* $\epsilon > 0$ *and* $\mathbf{u}, \mathbf{w} \in \mathbf{H}^1(\Omega)$ *and* $\mathbf{v} \in \mathbf{L}^\infty(\Omega) \cap \mathbf{H}^1(\Omega)$, *there exists a positive constant* C_ϵ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\mathbf{w}\|_1^2 + C_\epsilon \|\mathbf{u}\|_1^2 [\|\nabla \mathbf{v}\|^2 + \|\mathbf{v}\|_\infty^2]. \tag{A.2}$$

(iii) *For any* $\epsilon_1 > 0, \epsilon_2 > 0$ *and* $\mathbf{u}, \mathbf{w} \in \mathbf{H}^1(\Omega)$, $\mathbf{v} \in \mathbf{L}^\infty(\Omega) \cap \mathbf{H}^1(\Omega)$ *there exists a positive constants* C_{ϵ_1} *and* C_{ϵ_2} *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon_1 \|\mathbf{u}\|_1^2 + \epsilon_2 \|\mathbf{w}\|_1^2 + C_{\epsilon_1} \|\mathbf{w}\|^2 + C_{\epsilon_2} \|\mathbf{u}\|^2 [\|\mathbf{v}\|_\infty^2 + \|\nabla \mathbf{v}\|^4]. \tag{A.3}$$

(iv) *For any* $\epsilon > 0$, $\mathbf{u} \in \mathbf{H}^2(\Omega)$, $\mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega)$, *there exists a positive constants* $C_{1,\epsilon}$ *and* $C_{2,\epsilon}$ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\mathbf{v}\|_1^2 + C_{1,\epsilon} \|\mathbf{u}\|_2^2 [\|\mathbf{w}\|^2 + \|\mathbf{v}\|^2] + C_{2,\epsilon} \|\nabla \mathbf{w}\|^2. \tag{A.4}$$

(v) *For any* ϵ *and* $\mathbf{u} \in \mathbf{H}^1(\Omega)$ *and* $\mathbf{v} \in \mathbf{H}^2(\Omega)$, *there exists a positive constant* C_ϵ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{u})| \leq \epsilon \|\mathbf{u}\|_1^2 + C_\epsilon [\|\nabla \mathbf{v}\|_{L^4}^2 + \|\mathbf{v}\|_2^2] \|\mathbf{u}\|^2. \tag{A.5}$$

(vi) *For any* ϵ *and* $\mathbf{u}, \mathbf{w} \in \mathbf{H}^1(\Omega)$ *and* $\mathbf{v} \in \mathbf{H}^2(\Omega)$, *there exists a positive constant* C_ϵ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\mathbf{u}\|_1^2 + C_\epsilon [\|\nabla \mathbf{v}\|_{L^4}^2 + \|\mathbf{v}\|_2^2] \|\mathbf{w}\|_1^2. \tag{A.6}$$

(vii) *For any* $\epsilon_1 > 0, \epsilon_2 > 0$ *and* $\mathbf{u}, \mathbf{v} \in \mathbf{H}^1(\Omega)$, $\mathbf{w} \in \mathbf{L}^\infty(\Omega) \cap \mathbf{H}^1(\Omega)$ *there exists a positive constants* C_{ϵ_1} *and* C_{ϵ_2} *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon_1 \|\mathbf{u}\|_1^2 + \epsilon_2 \|\mathbf{v}\|_1^2 + C_{\epsilon_1} \|\mathbf{v}\|^2 + C_{\epsilon_2} \|\mathbf{u}\|^2 [\|\mathbf{w}\|_\infty^2 + \|\nabla \mathbf{w}\|^2]. \tag{A.7}$$

(viii) *For any* $\epsilon > 0$ *and* $\mathbf{u}, \mathbf{v} \in \mathbf{V}$, *there exists a positive constant* C_ϵ *such that*

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{u})| \leq \epsilon \|\nabla \mathbf{u}\|^2 + C_\epsilon \|\mathbf{u}\|^2 [\|\nabla \mathbf{v}\|^2 + \|\mathbf{v}\|_{L^4(\Omega)}^4]. \tag{A.8}$$

Proof.

(i) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|_\infty [\|\nabla \mathbf{v}\| \|\mathbf{w}\| + \|\mathbf{v}\| \|\nabla \mathbf{w}\|].$$

Since $\|\mathbf{u}\|_\infty \leq C \|\mathbf{u}\|_2$, the desired result follow by Young's inequality.

(ii) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{v}\| \|\mathbf{w}\|_{L^4} + C \|\mathbf{u}\| \|\nabla \mathbf{w}\| \|\mathbf{v}\|_\infty.$$

The result now follows by Sobolev embedding and Young's inequality.

(iii) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{v}\| \|\mathbf{w}\|_{L^4} + C \|\mathbf{u}\| \|\nabla \mathbf{w}\| \|\mathbf{v}\|_\infty.$$

The result now follows by Gagliardo−Nirenberg's inequality with $q = 2$ and $r = 4$ and Young's inequality.

(iv) The proof of this result similar to (i) and thus omitted.

(v) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{u})| \leq C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{v}\| \|\mathbf{u}\|_{L^4} + C \|\mathbf{u}\| \|\nabla \mathbf{u}\| \|\mathbf{v}\|_\infty.$$

By using Gagliardo−Nirenberg inequality, we obtain

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{u})| \leq C \|\mathbf{u}\| \|\mathbf{u}\|_1 [\|\nabla \mathbf{v}\| + \|\mathbf{v}\|_\infty].$$

The result now follows by Young's inequality.

(vi) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{v}\| \|\mathbf{w}\|_{L^4} + C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{w}\| \|\mathbf{v}\|_{L^4}.$$

(vii) By Hölder's inequality, we have

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|_{L^4} \|\nabla \mathbf{v}\| \|\mathbf{w}\|_{L^4} + C \|\mathbf{u}\| \|\nabla \mathbf{v}\| \|\mathbf{w}\|_\infty.$$

But by Gagliardo−Nirenberg's inequality with $q = 2$ and $r = 4$, we obtain

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq C \|\mathbf{u}\|^{\frac{1}{2}} \|\mathbf{u}\|_1^{\frac{1}{2}} \|\nabla \mathbf{w}\| \|\mathbf{v}\|^{\frac{1}{2}} \|\mathbf{v}\|_1^{\frac{1}{2}} + C \|\mathbf{u}\| \|\nabla \mathbf{v}\| \|\mathbf{w}\|_\infty.$$

Applying Young's inequality repeatedly, we get

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon_1 \|\mathbf{u}\|_1^2 + \epsilon_2 \|\mathbf{v}\|_1^2 + C \|\mathbf{u}\|^2 [\|\mathbf{w}\|_\infty^2 + \|\nabla \mathbf{w}\|^2] + C \|\mathbf{v}\|^2.$$

(viii) By Hölder's and Gagliardo−Nirenberg's inequalities, we obtain

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{u})| \leq C \|\mathbf{u}\| \|\nabla \mathbf{u}\| \|\nabla \mathbf{v}\| + \|\mathbf{u}\|^{\frac{1}{2}} \|\nabla \mathbf{u}\|^{\frac{3}{2}} \|\mathbf{v}\|_{L^4(\Omega)}.$$

Now the result follows by Young's inequality. □

References

[1] B.F. Armaly, F. Durst and J.C.F. Pereira, Experimental and theoretical investigation of backward-facing step flow. *J. Fluid Mech.* **127** (1983) 473–496.

[2] F. Ben Belgacem, H. El Fekih and J.P. Raymond, A penalized approach for solving a parabolic equation with non-smooth Dirichlet boundary conditions *Asymptot. Anal.* **34** (2003) 121–136.

[3] F. Ben Belgacem, C. Bernardi and H. El Fekih, Dirichlet boundary control for a parabolic equation with final observation: A space-time mixed formulation and penalization. *Asymptot. Anal.* **71** (2011) 101–121.

[4] E. Casas, M. Mateos and J.P. Raymond, Penalization of Dirichlet optimal control problems. *ESAIM: COCV* **15** (2009) 782–809.

[5] H.O. Fattorini and S.S. Sritharan, Existence of optimal controls for viscous flow problems. *Proc. R. Soc. Lond. Ser. A* **439** (1992) 81–102.

[6] A. Fursikov, M.D. Gunzburger and L.S. Hou, Boundary value problems and optimal boundary control for the Navier−Stokes system: the two-dimensional case. *SIAM J. Control Optim.* **36** (1998) 852–894

[7] E. Gagliardo, Proprieta di alcune classi di funzioni in piu variabili. *Ric. Mat.* **7** (1958) 102–137.

[8] M. Gad-el-Hak, A. Pollard and J.P. Bonnet, (eds.) Flow Control, Fundamentals and Practices. *Lecture Notes in Physics.* Springer, Berlin (1998).

[9] D.K. Gartling, A test problem for outfow boundary conditions-fow over a backward facing step. *Int. J. Numer. Methods Fluids* **11** (1990) 953–967.

[10] V. Girault and P.A. Raviart, *Finite Element Method for Navier−Stokes Equations.* Springer, Berlin (1986).

[11] M.D. Gunzburger, (ed.) Flow Control, IMA 68. Springer-Verlag, New York (1995).

[12] M.D. Gunzburger and L.S. Hou, Treating inhomogeneous essential boundary conditions in finite element and calculation of the boundary stresses. *SIAM J. Numer. Anal.* **29** (1992) 390–424.

[13] M.D. Gunzburger, L.S. Hou and Th.P. Svobodny, Analysis and finite approximation of optimal control problems for the stationary Navier−Stokes equations with Dirichlet control. *Math. Model. Numer. Anal.* **25** (1990) 711–748.

[14] J.G. Heywood and R. Rannacher, Finite element approximation of the non-stationary Navier−Stokes problem. Part 1: Regularity of solutions and second-order error estimates for spatial discretization. *SIAM J. Numer. Anal.* **19** (1982) 275–311.

[15] L.S. Hou and S.S. Ravindran, A Penalized Neumann Control Approach for Solving an Optimal Dirichlet Control Problem for the Navier−Stokes Equations. *SIAM J. Control Optim.* **36** (1998) 1795–1814.

[16] L.S. Hou and S.S. Ravindran, Numerical approximation of optimal flow control problems by a penalty method: Error estimates and numerical results. *SIAM J. Sci. Comput.* **20** (1999) 1753–1777.

[17] G. Hellwig, *Differential Operators of Mathematical Physics: An Introduction.* Addison-Wesley, Reading, MA (1967).

[18] J. Necas, *Les Méthods Directes en Théorie des Équations Elliptiques.* Masson et Cie, Paris (1967).

[19] L. Nirenberg, On elliptic partial differential equations. *Ann. Sci. Norm. Sup. Pisa* **13** (1959) 116–162.

[20] S.S. Ravindran, Reduced-Order Adaptive Controllers for Fluid Flows Using POD. *J. Sci. Comput.* **15** (2000) 457–478.

[21] S.S. Ravindran, Dirichlet Control of Unsteady Navier−Stokes type System related to Soret Convection by Boundary Penalty Method, *ESAIM: COCV* **40** (2014) 840–863.

[22] S.S. Ravindran, Convergence of Extrapolated BDF2 Finite Element Schemes For Unsteady Penetrative Convection Model. *Numer. Funct. Anal. Optim.* **33** (2012) 48–79.

[23] S.S. Sritharan, *Optimal Control of Viscous Flows.* SIAM, Philadelphia (1998).