Nonlinear Analysis 74 (2011) 585–599

Contents lists available at ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

Optimal control for multi-phase fluid Stokes problems

Karl Kunisch [a,*], Xiliang Lu [b,1]

[a] Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36 A-8010 Graz, Austria

[b] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria

ARTICLE INFO

Article history:
Received 4 November 2009
Accepted 5 September 2010

MSC:
49K20
76T99

Keywords:
Multi-phase fluids
Optimal control
Optimality system
Artificial diffusion
Renormalized solution

ABSTRACT

Optimal control for a system consistent of the viscosity dependent Stokes equations coupled with a transport equation for the viscosity is studied. Motivated by a lack of sufficient regularity of the adjoint equations, artificial diffusion is introduced to the transport equation. The asymptotic behavior of the regularized system is investigated. Optimality conditions for the regularized optimal control problems are obtained and again the asymptotic behavior is analyzed. The lack of uniqueness of solutions to the underlying system is another source of difficulties for the problem under investigation.

1. Introduction

The focus of this work is to establish an approach for optimal control multi-phase fluid flow. More specifically we consider the problem

$$\min J(\eta, \mathbf{u}) = \frac{1}{2}\|\eta - \tilde{\eta}\|^2_{L^2(Q)} + \frac{\alpha}{2}\|\mathcal{B}\mathbf{u}\|^2_{L^2(Q)}, \tag{1.1}$$

subject to

$$\begin{cases} \mathbf{y}_t - \operatorname{div}\,(\eta(\nabla \mathbf{y})) + \nabla p = \mathcal{B}\mathbf{u}, \\ \operatorname{div}\mathbf{y} = 0, \qquad \mathbf{y}|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{y}|_{t=0} = \mathbf{y}_0, \\ \eta_t + \mathbf{y}\cdot\nabla\eta = 0, \\ \eta|_{t=0} = \eta_0. \end{cases} \tag{1.2}$$

Let us describe the various terms in this problem formulation. Here Ω is a bounded domain in $\mathbb{R}^2$ with smooth boundary, time interval $T > 0$ is fixed and $Q = (0, T) \times \Omega$. The spatio-temperally dependent vector field $\mathbf{y}$ presents the velocity of the fluid, p its pressure, and η is the nonconstant viscosity of the fluid. Further $\mathbf{y}_0$ and η_0 are the initial velocity and viscosity respectively. The control variable is denoted by $\mathbf{u}$, it may act on the subset $\tilde{\Omega} \subset \Omega$. Control operator $\mathcal{B}$ is a bounded linear operator from $L^2(\mathbf{L}^2(\tilde{\Omega}))$ to $\mathbf{L}^2(Q)$, which will be defined in a later section. The control problem consists in finding $\mathbf{u}$ such that the corresponding state-control vector $(\mathbf{y}, \eta, p, \mathbf{u})$ minimizes $J(\eta, \mathbf{u})$, where $\tilde{\eta}$ is given and fixed.

* Corresponding author.
E-mail addresses: karl.kunisch@uni-graz.at (K. Kunisch), xiliang.lu@ricam.oeaw.ac.at (X. Lu).
1 Current address: Department of Mathematics, Wuhan University, China.

0362-546X/$ – see front matter Crown Copyright © 2010 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.na.2010.09.014

If $\tilde{\eta}$ is chosen as $\eta_1 + (\eta_2 - \eta_1)\chi_{\hat{\Omega}}$ where $\chi_{\hat{\Omega}}$ is the characteristic function of a set $\hat{\Omega} \subset \Omega$, then (1.1) and (1.2) represents the problem of determining a control $\mathbf{u}$ such that the interface between the two fluids with viscosities η_1 and η_2 close to the boundary $\partial\hat{\Omega}$ of $\hat{\Omega}$.

One of the key issues in optimal control with partial differential equations as constraints consists in establishing existence and first order necessary optimality conditions, which are expressed in the form of optimality systems. Here we shall establish existence by means of a compensated compactness argument. To obtain an optimality system, one can rely on a Lagrangian formalism, for example. Proceeding formally we introduce adjoint variables for the velocity, the pressure and the viscosity and denote them by $(\mathbf{z}, q, \xi)$, and denote by $e_i(\mathbf{y}, p, \eta) = 0, i = 1, 2, 3$ the momentum, the mass, and the transport equations respectively. Defining the formal Lagrangian

$$\mathcal{L}(\mathbf{y}, p, \eta, \mathbf{u}; \mathbf{z}, q, \xi) = J(\eta, \mathbf{u}) + \langle \mathbf{z}, e_1(\mathbf{y}, p, \eta)\rangle + \langle q, e_2(\mathbf{y}, p, \eta)\rangle + \langle \xi, e_3(\mathbf{y}, p, \eta)\rangle,$$

and setting the first derivatives with respect to $(\mathbf{y}, p, \eta, \mathbf{u})$ equal to zero, we obtain formal adjoint equations

$$\begin{cases} -\mathbf{z}_t - \operatorname{div}(\eta\nabla\mathbf{z}) - \eta\nabla\xi + \nabla q = \mathbf{0}, \\ \operatorname{div}\mathbf{z} = 0, \qquad \mathbf{z}|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{z}|_{t=T} = \mathbf{0}, \\ -\xi_t + \mathbf{y}\cdot\nabla\xi + \nabla\mathbf{y} : \nabla\mathbf{z} = -(\eta - \tilde{\eta}), \\ \xi|_{t=T} = 0, \end{cases} \tag{1.3}$$

where $\nabla\mathbf{y} : \nabla\mathbf{z}$ is the matrix inner product of Frobenius type, and the optimality condition:

$$\alpha\mathcal{B}^*\mathcal{B}\mathbf{u} = \mathcal{B}^*\mathbf{z}. \tag{1.4}$$

Combining the primal equations (1.2), the adjoint equations (1.3) and the optimality condition (1.4) provides the formal optimality system. These equations are indeed only formal since the transport equations in (1.2) and (1.3) have no smoothing properties. Hence $\mathbf{z}$ is strictly less regular in space than H^1 and ξ is strictly less regular in space than L^1. The bilinear coupling in (1.3) is the source of significant difficulties in analyzing this equation.

This lack of regularity of the adjoint equations motivates the introduction of a smoothing step. In the present work, we introduce artificial diffusion to the transport equation, which results in

$$\eta^\epsilon_t - \epsilon\triangle\eta^\epsilon + \mathbf{y}^\epsilon\cdot\nabla\eta^\epsilon = 0, \tag{1.5}$$

and investigate the optimal control problems for the regularized system.

Let us briefly outline the contents of the paper. In Section 2 we gather technical results which will be used throughout the remainder of the paper. The experienced reader can proceed directly to Section 3 where the regularized primal problems are investigated. The existence of solutions for each $\epsilon > 0$ is shown by means of Schauder's fixed point theorem. It is further shown that as $\epsilon \to 0^+$ limit points of regularized solutions satisfy (1.2), where the solution concept for the transport equations is that of regularized solutions in the sense of DiPerna–Lions (cf. [1]). In Section 4 the optimal control formulation is studied. An optimality system is derived and convergence of the optimal control problems as $\epsilon \to 0^+$ is investigated. The lack of uniqueness of solutions to (1.1) significantly complicates this analysis.

The investigations of this paper can certainly be extended in several aspects. Similar results as presented here should also hold true if the Stokes equations are replaced by the Navier Stokes equations with the nonconstant density function. More involved cost-functionals, and cost functionals involving the velocity $\mathbf{y}$ can be treated by the same techniques as used in this paper.

Finally let us give only a few comments on the multi-phase fluid model that is used in this paper. If $\eta_0 \in \{\eta_1, \ldots, \eta_L\}$, with η_i constants strictly larger than zero, then the renormalized solution $\eta(t, x) \in \{\eta_1, \ldots, \eta_L\}$ as well, see e.g. [2,3]. The transport equation in (1.2) is a variational formulation, posed on all of the domain Ω, of the immiscibility condition along the interfaces occupied by fluids with different viscosities, as proved in [2], Lemma 2.3. Of course, once the regularization is introduced the solution η_ϵ will not satisfy $\eta(t, x) \in \{\eta_1, \ldots, \eta_L\}$, but rather mushy regions will arise. In [4] an improved model is investigated, which allows for shear rate dependent viscosities and which takes into consideration surface tension along the interfaces of different fluids. A different analytic framework for (1.2) is based on viscosity solutions. Global existence is shown in [5] under the assumption that the difference of the viscosities of two different fluids is sufficiently small. Finally global existence to multiphase viscous flow is also investigated in [6], again under the condition that the viscosities of the fluids in different phases do not differ too much.

2. Preliminaries and notations

Let Ω be an open bounded domain in R^2 with Lipschitz boundary. We use standard notation $W^{m,p}$ and H^m for the Sobolev space, and we simplify the norm of H^m as $\|f\|_m = \|f\|_{H^m}$.

We will repeatedly use the following inequalities. The generic constant C does not depends on the choice of u.

- Poincaré inequality: For any $u \in H^1_0$ or $u \in H^1 \cap L^2_0$, we have

$$\|u\| \le C\|\nabla u\|.$$

- Hölder inequality:

$$\int_\Omega |fgh| \leq \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}$$

where $p, q, r > 1$ and $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$.
- Sobolev inequality:

$$\|u\|_{L^4} \leq C\|u\|^{1/2}\|u\|_1^{1/2},$$
$$\|u\|_{L^p} \leq C\|u\|_1, \quad \text{for any } 1 \leq p < \infty,$$

where $\Omega \subset R^2$;
- Gronwall's inequality:
Let $y(s)$ be a nonnegative, absolutely continuous function in $[0, s]$ and satisfy for almost every s, the differential inequality:

$$y'(s) \leq a(s)y(s) + b(s), \tag{2.1}$$

where $a(s)$ and $b(s)$ are nonnegative, summable functions in $[0, s]$. Then we have:

$$y(s) \leq e^{\int_0^s a(\tau)d\tau}\left[y(0) + \int_0^s b(\tau)d\tau\right]. \tag{2.2}$$

- Aubin's Lemma: (cf. [7])
Let X_0, X_1, X_2 be Banach spaces such that

$$X_0 \subset X_1 \subset X_2, \quad X_i \text{ is reflexive for } i = 0, 1,$$

and the injection of X_0 into X_1 is compact. Let $1 < p_i < \infty$, $i = 0, 1$. Then the space

$$W = L^{p_0}(X_0) \cap W^{1,p_1}(X_2)$$

is compactly imbedded in $L^{p_0}(X_1)$.

For the Stokes equation, the following divergence free spaces are useful.

$$\mathcal{V} = \{\mathbf{u} \in C_0^\infty(\Omega), \text{div } \mathbf{u} = 0\},$$
$$\mathbf{H} = \overline{\mathcal{V}}^{\mathbf{L}^2} = \{\mathbf{u} \in \mathbf{L}^2, \text{div } \mathbf{u} = 0, \mathbf{u}\cdot\mathbf{n} = 0\},$$
$$\mathbf{V} = \overline{\mathcal{V}}^{\mathbf{H}_0^1} = \{\mathbf{u} \in \mathbf{H}_0^1, \text{div } \mathbf{u} = 0\},$$

where $\mathbf{H}$ and $\mathbf{V}$ are equipped with the norm induced by $\mathbf{L}^2$ and $\mathbf{H}_0^1$. We identify the dual space of $\mathbf{H}$ as itself, and define the dual space of $\mathbf{V}$ as $\mathbf{V}^*$. We also introduce the projection operator P from $\mathbf{L}^2$ to its divergence free subspace $\mathbf{H}$. By the Helmholtz–Hodge decomposition theorem (cf. [8]), we have:

$$\mathbf{L}^2 = \mathbf{H} \oplus \nabla H^1.$$

Now we introduce time dependent function spaces. For any function $f : [0, T] \to B$, where B is a given Banach space, we denote $f \in L^p(0, T; B)$ if

$$\begin{cases} \int_0^T \|f(t)\|_B^p dt < \infty, & 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{0\leq t\leq T} \|f(t)\|_B < \infty, & p = \infty. \end{cases}$$

The norm is defined by $\|f\|_{L^p(B)} = (\int_0^T \|f(t)\|_B^p)^{1/p}$ for $1 \leq p < \infty$ and $\|f\|_{L^\infty(B)} = \operatorname{ess\,sup}_{0\leq t\leq T} \|f(t)\|_B$. If $f(0) \in B$ and $\frac{\partial f}{\partial t} \in L^2(B)$, we denote it by $f \in H^1(0, T; B)$. We will use Q to represent the time-space domain, i.e.

$$Q = (0, T) \times \Omega.$$

For time dependent test functions, we denote:

$$\mathcal{V}_T = C^1([0, T], \mathcal{V}), \qquad C_T = C^1([0, T], C_0^\infty(\Omega)).$$

For any function $f \in C_T$ or $\mathbf{f} \in \mathcal{V}_T$, we can define an operator γ_0 by taking the initial data: $\gamma_0(f) = f(0)$ or $\gamma_0(\mathbf{f}) = \mathbf{f}(0)$. We will use this operator in Section 4.

The Stokes operator is defined as $\Lambda : \mathbf{V} \to \mathbf{V}^*$, $(\nabla\mathbf{y}, \nabla\mathbf{v}) = \langle\Lambda\mathbf{y}, \mathbf{v}\rangle_{\mathbf{V}^*,\mathbf{V}}$, for all $\mathbf{v} \in \mathbf{V}$. Since the domain is regular, the Stokes operator Λ can also be defined from $\mathbf{H}^2 \cap \mathbf{V}$ to $\mathbf{H}$ by $\Lambda\mathbf{y} = -P\triangle\mathbf{y}$. We will not distinguish the above two operators and denote them both by Λ. One can verify the Stokes operator has the following properties (see [8]):

- Λ is positive, self-adjoint operator,
- Λ is bijective from $\mathbf{V}$ to $\mathbf{V}^*$, hence it is an isomorphism from $\mathbf{V}$ to $\mathbf{V}^*$,

- $\|\mathbf{y}\|_1 \approx \|\Lambda^{1/2}\mathbf{y}\| \approx \|\Lambda\mathbf{y}\|_{\mathbf{V}^*}$, for all $\mathbf{y} \in \mathbf{V}$,
- $\|\mathbf{u}\|_{\mathbf{V}^*} \approx \|\Lambda^{-1/2}\mathbf{u}\| \approx \|\Lambda^{-1}\mathbf{u}\|_1$, for all $\mathbf{u} \in \mathbf{V}^*$,
- Λ is bijective from $\mathbf{H}^2 \cap \mathbf{V}$ to $\mathbf{H}$, hence it is an isomorphism from $\mathbf{H}^2 \cap \mathbf{V}$ to $\mathbf{H}$,
- $\|\mathbf{y}\|_2 \approx \|\Lambda^{1/2}\mathbf{y}\|_1 \approx \|\Lambda\mathbf{y}\|$, for all $\mathbf{y} \in \mathbf{H}^2 \cap \mathbf{V}$,
- $\|\mathbf{u}\| \approx \|\Lambda^{-1/2}\mathbf{u}\|_1 \approx \|\Lambda^{-1}\mathbf{u}\|_2$, for all $\mathbf{u} \in \mathbf{H}$,

where $\approx$ means that the norms are equivalent.

The control operator $\mathcal{B} \in \mathcal{L}(L^2(\mathbf{L}^2(\tilde{\Omega})), \mathbf{L}^2(Q))$, denotes the extension-by-zero operator into the domain $\Omega \setminus \tilde{\Omega}$. Hence $\mathcal{B}$ is an isometry from $L^2(\mathbf{L}^2(\tilde{\Omega}))$ to a subspace of $\mathbf{L}^2(Q)$, i.e.

$$\|\mathbf{u}\|_{L^2(L^2(\tilde{\Omega}))} = \|\mathcal{B}\mathbf{u}\|_{L^2(Q)}.$$

We denote two positive constants m and M as the lower bound and upper bound of viscosity, and the initial viscosity η_0 satisfies

$$0 < m \le \eta_0(\mathbf{x}) \le M < \infty.$$

The generic constants C and C_i only depend on Ω, m, M, T, the initial velocity $\mathbf{y}_0$, the initial viscosity η_0 and external force. C may be different in different cases, and C_i can be fixed in advance. C_ϵ is also constant but may depend on the choice of ϵ.

3. Existence and convergence for the approximated system

As we mentioned in the introduction, the governing equation for the multi-phases immiscible incompressible fluid reads:

$$\mathbf{y}_t - \text{div}\,(\eta\nabla\mathbf{y}) + \nabla p = \mathcal{B}\mathbf{u}, \tag{3.1}$$

$$\text{div}\,\mathbf{y} = 0, \tag{3.2}$$

$$\mathbf{y}|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{y}|_{t=0} = \mathbf{y}_0, \tag{3.3}$$

$$\eta_t + \mathbf{y}\cdot\nabla\eta = 0, \tag{3.4}$$

$$\eta|_{t=0} = \eta_0 \tag{3.5}$$

where $\mathbf{y}$ presents the velocity of the fluid, η is the viscosity of the fluid, $\mathbf{y}_0$ and η_0 are the initial velocity and viscosity respectively. To avoid the pressure term, we can put the system into the weak form: given $\mathcal{B}\mathbf{u} \in L^2(\mathbf{V}^*)$ and $\mathbf{y}_0 \in \mathbf{H}$, find $(\eta, \mathbf{y}) \in L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$, such that

$$\mathbf{y}_t - \text{div}\,(\eta\nabla\mathbf{y}) = \mathcal{B}\mathbf{u}, \quad \text{in } L^2(\mathbf{V}^*),$$

$$\eta_t + \mathbf{y}\cdot\nabla\eta = 0, \quad \text{in } L^2(H^{-1}),$$

with the initial conditions (3.3) and (3.5). The existence of a solution can be found in [2,3].

We take a singular perturbation to the system (3.1)–(3.5) and arrive at the following approximating system:

$$\mathbf{y}^\epsilon_t - \text{div}\,(\eta^\epsilon(\nabla\mathbf{y}^\epsilon)) + \nabla p^\epsilon = \mathcal{B}\mathbf{u}, \tag{3.6}$$

$$\text{div}\,\mathbf{y}^\epsilon = 0, \tag{3.7}$$

$$\mathbf{y}^\epsilon|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{y}^\epsilon|_{t=0} = \mathbf{y}_0, \tag{3.8}$$

$$\eta^\epsilon_t - \epsilon\triangle\eta^\epsilon + \mathbf{y}^\epsilon\cdot\nabla\eta^\epsilon = 0, \tag{3.9}$$

$$\eta^\epsilon|_{t=0} = \eta^\epsilon_0, \qquad \eta^\epsilon|_{\partial\Omega} = m, \tag{3.10}$$

where ϵ is a positive constant and η^ϵ_0 is an approximation of η_0 which satisfies

$$m \le \eta^\epsilon_0 \le M, \quad \text{a.e.}, \eta^\epsilon_0 \to \eta_0 \text{ in } L^2(\Omega) \text{ as } \epsilon \to 0^+. \tag{3.11}$$

To avoid the pressure term, we can consider the following equivalent approximating system. Given $\mathcal{B}\mathbf{u} \in L^2(\mathbf{V}^*)$ and $\mathbf{y}_0 \in \mathbf{H}$, find $(\eta^\epsilon, \mathbf{y}^\epsilon) \in L^2(\mathbf{H}^1) \cap L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap L^\infty(\mathbf{H}) \cap H^1(\mathbf{V}^*)$, such that

$$\mathbf{y}^\epsilon_t - \text{div}\,(\eta^\epsilon(\nabla\mathbf{y}^\epsilon)) = \mathcal{B}\mathbf{u}, \quad \text{in } L^2(\mathbf{V}^*), \tag{3.12}$$

$$\eta^\epsilon_t - \epsilon\triangle\eta^\epsilon + \mathbf{y}^\epsilon\cdot\nabla\eta^\epsilon = 0, \quad \text{in } L^2(H^{-1}), \tag{3.13}$$

with the initial conditions and boundary conditions (3.8) and (3.10). Moreover, if $\mathcal{B}\mathbf{u} \in \mathbf{L}^2(Q)$, $\mathbf{y}_0 \in \mathbf{V}$ and $\eta^\epsilon_0 - m \in H^1_0$, we find $(\eta^\epsilon, \mathbf{y}^\epsilon) \in L^2(\mathbf{H}^2) \cap L^\infty(Q) \cap L^\infty(H^1) \cap H^1(L^2) \times L^2(\mathbf{H}^2) \cap L^\infty(\mathbf{V}) \cap H^1(\mathbf{H})$, such that,

$$\mathbf{y}^\epsilon_t - P(\text{div}\,(\eta^\epsilon\nabla\mathbf{y}^\epsilon)) = P\mathcal{B}\mathbf{u}, \quad \text{in } L^2(\mathbf{H}), \tag{3.14}$$

$$\eta^\epsilon_t - \epsilon\triangle\eta^\epsilon + \mathbf{y}^\epsilon\cdot\nabla\eta^\epsilon = 0, \quad \text{in } L^2(Q), \tag{3.15}$$

with the initial conditions and boundary conditions (3.8) and (3.10). The equivalence of (3.6)–(3.10) to (3.12)–(3.13) and (3.14)–(3.15) is based on Helmholtz–Hodge decomposition theorem (cf. [8]). For simplicity, we will not distinguish the PDE form with its variational form in the paper, i.e. the pressure is omitted in the statement and the proof.

During this section, $\mathbf{u}$ is fixed as a given function. For simplicity we denote $\mathcal{B}\mathbf{u}$ by $\mathbf{f}$. Then we have the following existence and convergence results.

Theorem 3.1 (*Existence for Fixed* ϵ). *For any given positive constant* ϵ, $\mathbf{f} \in L^2(\mathbf{V}^*)$, $\mathbf{y}_0 \in \mathbf{H}$, $\eta_0^\epsilon \in L^2$ *with* $m \le \eta_0^\epsilon \le M$ *a.e., system (3.6)–(3.10) has at least one solution which satisfies* $(\eta^\epsilon, \mathbf{y}^\epsilon) \in L^2(\mathbf{H}^1) \cap L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap L^\infty(\mathbf{H}) \cap H^1(\mathbf{V}^*)$, *moreover* $m \le \eta^\epsilon \le M$.

The proof for existence is based on fixed point argument, and is given in Sections 3.1–3.3.

Theorem 3.2 (*Convergence for* $\epsilon \to 0^+$). *Assume that* $\mathbf{f} \in L^2(\mathbf{V}^*)$, $\mathbf{y}_0 \in \mathbf{H}$, *and* η_0^ϵ *satisfies (3.11) as* $\epsilon \to 0^+$. *Let* $(\eta^\epsilon, \mathbf{y}^\epsilon) \in L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$ *be any solution of (3.6)–(3.10). Then there exists a sequence* $\epsilon^n \to 0^+$ *such that* $\mathbf{y}^n \to \mathbf{y}$ *in* $\mathbf{L}^2(Q)$, $\eta^n \to \eta$ *in* $L^2(Q)$ *and* $(\eta, \mathbf{y}) \in L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$ *satisfies the Eqs. (3.1)–(3.5), with* $m \le \eta \le M$ *and* η *is a renormalized solution.*

The definition and property of renormalized solution can be found in [2]. The proof of convergence is given in Section 3.4.

3.1. A-priori estimate for the Stokes equation

We fix η as a measurable function satisfying $m \le \eta \le M$, a.e. For the Stokes equation

$$\mathbf{y}_t - \operatorname{div}(\eta \nabla \mathbf{y}) + \nabla p = \mathbf{f}, \tag{3.16}$$

$$\operatorname{div} \mathbf{y} = 0, \tag{3.17}$$

$$\mathbf{y}|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{y}|_{t=0} = \mathbf{y}_0, \tag{3.18}$$

existence and uniqueness of a weak solution can be obtained by standard argument similar to the constant viscosity case, and we have the *a*-priori estimate

$$\sup_{0\le t\le T} \|\mathbf{y}(t,\cdot)\|^2 + \int_0^t \|\mathbf{y}\|_1^2 \,\mathrm{d}t \le \|\mathbf{y}_0\| + C\int_0^T \|\mathbf{f}\|_{\mathbf{V}^*}^2 \mathrm{d}t. \tag{3.19}$$

Then $\mathbf{y}_t$ can be estimated as:

$$\|\mathbf{y}_t\|_{\mathbf{V}^*} = \sup_{\mathbf{v}\in\mathbf{V}} \frac{\langle \mathbf{y}_t, \mathbf{v}\rangle}{\|\nabla \mathbf{v}\|} = \sup_{\mathbf{v}\in\mathbf{V}} \frac{\langle \mathbf{f}, \mathbf{v}\rangle - (\eta\nabla\mathbf{y}, \nabla\mathbf{v})}{\|\nabla\mathbf{v}\|} \le \|\mathbf{f}\|_{\mathbf{V}^*} + M\|\nabla\mathbf{y}\|, \tag{3.20}$$

and by virtue of (3.19), we have

$$\|\mathbf{y}_t\|_{L^2(\mathbf{V}^*)} \le C\|\mathbf{f}\|_{L^2(\mathbf{V}^*)}. \tag{3.21}$$

We notice that the estimates (3.19) and (3.21) only depend on $\mathbf{f}$, $\mathbf{y}_0$, Ω, m and M. Hence there exists a constant C_1 such that:

$$\|\mathbf{y}\|_{L^\infty(\mathbf{H})} + \|\mathbf{y}\|_{L^2(\mathbf{V})} + \|\mathbf{y}_t\|_{L^2(\mathbf{V}^*)} \le C_1. \tag{3.22}$$

To avoid the pressure term in the Stokes equations (3.16)–(3.18) and make the proof simple, we can rewrite the Stokes equation in the following equivalent form: find $\mathbf{y} \in L^2(\mathbf{V}) \cap L^\infty(\mathbf{H}) \cap H^1(\mathbf{V}^*)$ which satisfies

$$\mathbf{y}_t - \operatorname{div}(\eta(\nabla\mathbf{y})) = \mathbf{f}, \quad \text{in } \mathcal{V}_T^*$$

$$\mathbf{y}|_{t=0} = \mathbf{y}_0.$$

3.2. A-priori estimate for the convection–diffusion equation

For $\epsilon > 0$ as a fixed positive constant, and given $\mathbf{y}$ which satisfies (3.22), we consider the following convection–diffusion equation:

$$\eta_t - \epsilon\triangle\eta + \mathbf{y}\cdot\nabla\eta = 0, \tag{3.23}$$

$$\eta|_{t=0} = \eta_0^\epsilon, \qquad \eta|_{\partial\Omega} = m. \tag{3.24}$$

Existence and uniqueness of the weak solution can be obtained by standard arguments, and we have the following a priori estimate:

$$\|\eta(t,\cdot) - m\|^2 + 2\epsilon\int_0^t \|\nabla\eta\|^2 = \|\eta_0^\epsilon - m\|^2. \tag{3.25}$$

In fact, shifting η by a constant function m, we find that $\bar\eta = \eta - m$ satisfies the same parabolic equation (3.23) with initial condition $\eta_0^\epsilon - m$ and homogenous Dirichlet boundary condition. Multiplying the resulting equation (3.23) by $\bar\eta$ gives (3.25).

By virtue the maximum principle for parabolic equations, we have

$$m \le \eta(t, \mathbf{x}) \le M. \tag{3.26}$$

The maximum principle we used here is Theorem 7.2 in [9] pp. 188. We need to check that the coefficient $\mathbf{y} = (y_1, y_2)$ satisfies $y_i^2 \in L^2(Q)$, i.e., $r = q = 2$ in that theorem. From Sobolev's inequality, we have

$$\int_0^T \int_\Omega y_i^4 \mathrm{d}x\mathrm{d}t \le \int_0^T \|\mathbf{y}\|_{L^4}^4 \mathrm{d}t \le C \int_0^T \|\mathbf{y}\|^2 \|\mathbf{y}\|_1^2 \mathrm{d}t,$$

and since $\mathbf{y} \in L^2(\mathbf{V}) \cap L^\infty(\mathbf{H})$, we obtain $y_i^2 \in L^2(Q)$.

We proceed to estimate the time derivatives:

$$\begin{aligned}\|\eta_t\|_{H^{-1}} &= \sup_{v \in H_0^1} \frac{\langle \eta_t, v\rangle}{\|\nabla v\|} = \sup_{v \in H_0^1} \frac{\langle \epsilon \triangle \eta - \mathbf{y} \cdot \nabla \eta, v\rangle}{\|\nabla v\|} \\ &= \sup_{v \in H_0^1} \frac{-\epsilon(\nabla \eta, \nabla v) + (\mathbf{y}\eta, \nabla v)}{\|\nabla v\|} = \|\mathbf{y}\eta - \epsilon \nabla \eta\|.\end{aligned}$$

With the help of (3.25), (3.26) and since $\mathbf{y} \in L^2(\mathbf{V})$, we obtain

$$\|\eta\|_{L^2(H^{-1})} \le C. \tag{3.27}$$

Since the constants in (3.25) and (3.27) only depend on m, M and C_1, we can define two constants C_2 and $C_{2,\epsilon}$ such that

$$\|\eta\|_{L^2(Q)} + \|\eta_t\|_{L^2(H^{-1})} \le C_2, \tag{3.28}$$

$$\|\eta\|_{L^2(H^1)} + \|\eta_t\|_{L^2(H^{-1})} \le C_{2,\epsilon}, \tag{3.29}$$

Similar to the Stokes equation, we can also rewrite Eqs. (3.23)–(3.24) in the following equivalent way: find $\eta \in L^2(H^1) \cap H^1(H^{-1})$, such that

$$\begin{aligned}&\eta_t - \epsilon \triangle \eta + \mathbf{y} \cdot \nabla \eta = 0, \quad \text{in } C_T^*, \\ &\eta|_{t=0} = \eta_0, \qquad \eta|_{\partial\Omega} = m.\end{aligned}$$

3.3. *Proof of Theorem 3.1*

We will prove existence for the approximating system (3.6)–(3.10) by Schauder's fixed point theorem (cf. [10]). Since ϵ is fixed in Theorem 3.1, the notation η and $\mathbf{y}$ without subscript ϵ are used in the proof for simplicity of notation. We define two Banach spaces as

$$E_1 = L^2(H^1) \cap H^1(H^{-1}), \qquad \mathbf{E}_2 = L^\infty(\mathbf{H}) \cap L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*). \tag{3.30}$$

Let $K_1 \subset E_1$ and $\mathbf{K}_2 \subset \mathbf{E}_2$ be given by:

$$\begin{aligned}&K_1 = \{\eta : \|\eta\|_{E_1} \le C_{2,\epsilon},\ m \le \eta \le M \text{ a.e.}\} \\ &\mathbf{K}_2 = \{\mathbf{y} : \|\mathbf{y}\|_{\mathbf{E}_2} \le C_1\},\end{aligned}$$

where C_1 and $C_{2,\epsilon}$ are defined in estimates (3.22) and (3.29). For any $\eta \in K_1$, the estimate (3.22) guarantees the existence of a velocity $\mathbf{y} \in \mathbf{K}_2$ to the Stokes equations (3.16)–(3.18). Similarly for any $\mathbf{y} \in \mathbf{K}_2$, the estimates (3.26) and (3.29) imply that the convection–diffusion equations (3.23)–(3.24) can be solved for a new viscosity $\xi \in K_1$. This combined mapping is defined as $\tau(\eta) = \xi$ and we note that τ maps K_1 into itself. Theorem 3.1 follows directly from the following lemma.

Lemma 3.3. *Let the same assumption hold in Theorem 3.1. Then the map τ defined as above is well-defined and has at least one fixed point in K_1.*

Proof. Clearly τ is well-defined. We will prove existence for the fixed point. First we notice that K_1 is a closed, bounded and convex set in E_1 and $\mathbf{K}_2$ is also a closed bounded set in $\mathbf{E}_2$. To apply Schauder's fixed point theorem, we need to prove that τ is continuous and compact. While proving these facts we always use the notation $\eta \to \mathbf{y} \to \xi$ (with or without subscript n). The proof is based on the following claims:

1. If $\{\eta^n\} \subset K_1$ is a weak convergent sequence in E_1 with limit η, then there exists a subsequence $\{\mathbf{y}^{n_j}\} \subset \mathbf{K}_2$ and $\mathbf{y} \in \mathbf{K}_2$ such that $\mathbf{y}^{n_j} \rightharpoonup \mathbf{y}$ weak-star in $\mathbf{E}_2$.
2. If $\{\mathbf{y}^n\} \subset \mathbf{K}_2$ is a weak star convergent sequence in $\mathbf{E}_2$ with limit $\mathbf{y}$, then there exists a subsequence $\{\xi^{n_j}\} \subset K_1$ and $\xi \in K_1$ such that $\xi^{n_j} \to \xi$ strongly in E_1.

3. Consider a sequence $\{a^n\}$ in a Banach space B. If for any subsequence of $\{a^n\}$ (denoted by $\{a^{n_i}\}$), we can pick up a sub-subsequence $\{a^{n_{i_j}}\}$ such that $\{a^{n_{i_j}}\}$ converges to $a \in B$ in the strong or the weak or the weak-star topology, then $\{a^n\}$ converges to a in the same topology.

Proof of claim 1. Since $m \le \eta^n \le M$, a.e., estimate (3.22) implies that $\{\mathbf{y}^n\} \subset \mathbf{K}_2$. Hence we can choose a subsequence $\{\mathbf{y}^{n_i}\}$ (still denoted by $\{\mathbf{y}^n\}$ for simplicity) and $\mathbf{z} \in \mathbf{K}_2$, such that

$$\begin{aligned} &\mathbf{y}^n \rightharpoonup \mathbf{z} \quad \text{in } L^2(\mathbf{V}),\\ &\mathbf{y}^n \rightharpoonup \mathbf{z} \quad \text{in } L^\infty(\mathbf{H}) \text{ in the weak-star topology},\\ &\mathbf{y}^n_t \rightharpoonup \mathbf{z}_t \quad \text{in } L^2(\mathbf{V}^*). \end{aligned}$$

Since $\{\eta^n\} \subset K_1$, Aubin's lemma implies that there exists a subsequence (still denoted by η^n) which converges to η strongly in $L^2(Q)$. Then strong convergence of η^n in $L^2(Q)$ and weak convergence of $\nabla\mathbf{y}^n$ in $L^2(Q)$ imply convergence of $\eta^n\nabla\mathbf{y}^n$ in the distribution sense. By definition, $\mathbf{y}^n$ solves the Stokes equation

$$\mathbf{y}^n_t - \operatorname{div}\,(\eta^n(\nabla\mathbf{y}^n)) = \mathbf{f} \quad \text{in } \mathcal{V}_T^*.$$

After passage to the limit, for any test function $\mathbf{v} \in \mathcal{V}_T$, we have

$$\begin{aligned} &\langle \mathbf{y}^n_t, \mathbf{v}\rangle_Q \to \langle \mathbf{z}_t, \mathbf{v}\rangle_Q,\\ &\langle \operatorname{div}\,(\eta^n(\nabla\mathbf{y}^n)), \mathbf{v}\rangle_Q = -\langle \eta^n(\nabla\mathbf{y}^n), \nabla\mathbf{v}\rangle_Q\\ &\quad \to -\langle \eta(\nabla\mathbf{z}), \nabla\mathbf{v}\rangle_Q = \langle \operatorname{div}\,(\eta(\nabla\mathbf{z})), \mathbf{v}\rangle_Q, \end{aligned}$$

and hence $\mathbf{z}$ satisfies the following equation

$$\mathbf{z}_t - \operatorname{div}\,(\eta(\nabla\mathbf{z})) = \mathbf{f} \quad \text{in } \mathcal{V}_T^*.$$

For the initial condition, we notice that $\mathbf{E}_2$ is compactly embedded into $C([0,T],\mathbf{V}^*)$ (see [11]). Since $\mathbf{y}^n|_{t=0} = \mathbf{y}_0$ for all n, we have $\mathbf{z}|_{t=0} = \mathbf{y}_0$. Then uniqueness of the Stokes equation implies that $\mathbf{z} = \mathbf{y}(\eta)$.

Proof of claim 2. Since $\{\mathbf{y}^n\} \subset \mathbf{K}_2$, inequality (3.29) implies that $\{\xi^n\} \subset K_1$. We can choose a subsequence (still denoted by $\{\xi^n\}$) and $\psi \in K_1$ such that

$$\begin{aligned} &\xi^n \rightharpoonup \psi \quad \text{in } L^2(H^1),\\ &\xi^n_t \rightharpoonup \psi_t \quad \text{in } L^2(H^{-1}). \end{aligned}$$

By definition, ξ^n satisfies the equation

$$\xi^n_t - \epsilon\triangle\xi^n + \mathbf{y}^n\cdot\nabla\xi^n = 0 \quad \text{in } \mathcal{C}_T^*.$$

Since $\mathbf{y}^n \in \mathbf{K}_2$, Aubin's Lemma implies that there exists a subsequence (still denoted by $\mathbf{y}^n$) converging to $\mathbf{y}$ strongly in $L^2(\mathbf{H})$. Then strong convergence of $\mathbf{y}^n$ in $L^2(Q)$ and weak convergence of $\nabla\xi^n$ in $L^2(Q)$ imply convergence of $\mathbf{y}^n\cdot\nabla\xi^n$ in the distribution sense. After passage to the limit, for any test function $\phi \in C_T$, we have

$$\begin{aligned} &\langle \xi^n_t, \phi\rangle_Q \to \langle \psi_t, \phi\rangle_Q,\\ &\langle \triangle\xi^n, \phi\rangle_Q \to \langle \triangle\psi, \phi\rangle_Q,\\ &\langle \mathbf{y}^n\cdot\nabla\xi^n, \phi\rangle_Q \to \langle \mathbf{y}\cdot\nabla\psi, \phi\rangle_Q. \end{aligned}$$

This implies that ψ satisfies the equation

$$\psi_t - \epsilon\triangle\psi + \mathbf{y}\cdot\nabla\psi = 0 \quad \text{in } \mathcal{C}_T^*.$$

The initial condition can be treated similarly as in the proof of claim 1. Since E_1 is compactly embedded into $C([0,T],H^{-1})$, we have $\psi|_{t=0} = \eta_0^\epsilon$. The boundary condition can also be treated by shifting every function by the constant function m and replacing H^1 by H_0^1 in the proof. Then by virtue of the uniqueness of the convection–diffusion equation, we have $\psi = \xi$.

Now we need to prove the strong convergence of $\{\xi^n\}$ in E_1. Defining $\varphi^n = \xi^n - \xi$, $\mathbf{z}^n = \mathbf{y}^n - \mathbf{y}$, we find that φ^n satisfies

$$\varphi^n_t - \epsilon\triangle\varphi^n + \mathbf{z}^n\cdot\nabla\xi^n + \mathbf{y}\cdot\nabla\varphi^n = 0,$$

with zero boundary and initial conditions. Multiplying φ^n on both sides of the above equation, we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\varphi^n\|^2 + \epsilon\|\nabla\varphi^n\|^2 + \langle \mathbf{z}^n\cdot\nabla\xi^n, \varphi^n\rangle = 0. \tag{3.31}$$

Since $\mathbf{z}^n$ is divergence free and $m \le \xi^n \le M$ a.e., we have

$$\langle \mathbf{z}^n \cdot \nabla \xi^n, \varphi^n \rangle = -\langle \mathbf{z}^n \xi^n, \nabla \varphi^n \rangle \le \frac{\epsilon}{2} \|\nabla \varphi^n\|^2 + C_\epsilon \|\mathbf{z}^n\|^2.$$

Substituting into (3.31) and integrating in time gives

$$\sup_{0\le t\le T} \|\varphi^n\|^2 + \int_0^T \|\varphi^n\|_1^2 \mathrm{d}t \le C_\epsilon \int_0^T \|\mathbf{z}^n\|^2 \mathrm{d}t. \tag{3.32}$$

The time derivative can be evaluated as

$$\begin{aligned}\|\varphi_t^n\|_{H^{-1}} &= \sup_{v\in H_0^1} \frac{\langle \varphi_t^n, v\rangle}{\|\nabla v\|} = \sup_{v\in H_0^1} \frac{\langle -\epsilon \nabla \varphi^n + \xi^n \mathbf{z}^n + \varphi^n \mathbf{y}^n, \nabla v\rangle}{\|\nabla v\|} \\ &\le \epsilon \|\varphi^n\|_1 + M\|\mathbf{z}^n\| + \|\mathbf{y}^n\|_{L^4}^{1/2} \|\varphi^n\|_{L^4}^{1/2} \le C_\epsilon (\|\varphi^n\|_1 + \|\mathbf{z}^n\| + \|\mathbf{y}^n\| \, \|\varphi^n\|_1 + \|\mathbf{y}^n\|_1 \|\varphi^n\|).\end{aligned}$$

Since $\mathbf{y}^n \in \mathbf{K}_2$, we have

$$\|\varphi_t^n\|_{L^2(H^{-1})} \le C_\epsilon (\|\varphi^n\|_{L^2(H^1)} + \|\mathbf{z}^n\|_{L^2(Q)} + \|\varphi^n\|_{L^\infty(L^2)}).$$

By virtue of $\|\mathbf{z}^n\|_{L^2(Q)} \to 0$, we conclude that $\varphi^n \to 0$ strongly in X.

Proof of claim 3. See [12].

Proof of the lemma. Continuity of τ. Consider a sequence $\{\eta^n\} \subset K_1$ and $\eta \in K_1$, such that $\eta^n \to \eta$ in E_1. By virtue of claims 1 and 3, we obtain $\{\mathbf{y}^n\} \subset \mathbf{K}_2$, $\mathbf{y} \in \mathbf{K}_2$ and $\mathbf{y}^n \rightharpoonup \mathbf{y}$ in $\mathbf{E}_2$ in the weak star topology. Then by claims 2 and 3, we have $\xi^n \to \xi$ in E_1 as desired. Compactness of τ follows from claims 1 and 2. □

3.4. *Proof of Theorem 3.2*

From Theorem 3.1, there exists at least one solution $(\eta^\epsilon, \mathbf{y}^\epsilon) \in K_1 \times \mathbf{K}_2$ for the system (3.6)–(3.10). From estimates (3.22), (3.25), (3.26) and (3.28), we have the *a*-priori estimates

$$m \le \eta^\epsilon \le M \quad \text{a.e.,} \tag{3.33}$$

$$\|\mathbf{y}^\epsilon\|_{L^\infty(\mathbf{H})} + \|\mathbf{y}^\epsilon\|_{L^2(\mathbf{V})} + \|\mathbf{y}_t^\epsilon\|_{L^2(\mathbf{V}^*)} \le C_1, \tag{3.34}$$

$$\|\eta^\epsilon(t,\cdot) - m\| \le \|\eta_0^\epsilon - m\|, \qquad \|\eta_t^\epsilon\|_{L^2(H^{-1})} \le C_2. \tag{3.35}$$

Since $\{(\eta^\epsilon, \mathbf{y}^\epsilon)\}$ is a bounded set in $L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$, we can find a subsequence which is denoted by $(\eta^n, \mathbf{y}^n)$ and $(\xi, \mathbf{z})$ in $L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$, such that

$$\begin{aligned}&\mathbf{y}^n \rightharpoonup \mathbf{z} \quad \text{in } L^2(\mathbf{V}),\\ &\mathbf{y}_t^n \rightharpoonup \mathbf{z}_t \quad \text{in } L^2(\mathbf{V}^*),\\ &\eta^n \rightharpoonup \xi \quad \text{in } L^\infty(Q) \text{ weak star},\\ &\eta_t^n \rightharpoonup \xi_t \quad \text{in } L^2(H^{-1}).\end{aligned}$$

We recall that $(\eta^n, \mathbf{y}^n)$ satisfy the equations:

$$\begin{aligned}&\mathbf{y}_t^n - \operatorname{div}\,(\eta^n(\nabla \mathbf{y}^n)) = \mathbf{f} \quad \text{in } \mathcal{V}_T^*,\\ &\eta_t^n - \epsilon^n \triangle \eta^n + \mathbf{y}^n \cdot \nabla \eta^n = 0, \quad \text{in } C_T^*,\end{aligned}$$

with the same initial and boundary condition as (3.8) and (3.10). Since $\mathbf{y}^n \to \mathbf{z}$ strongly in $L^2(Q)$ by Aubin's Lemma and $\eta^n \rightharpoonup \xi$ in $L^2(Q)$, we have $\mathbf{y}^n \eta^n$ converges to $\mathbf{z}\xi$ in the distribution sense. Choosing a test function $\phi \in C_T$ in the convection–diffusion equation, we find for $\epsilon^n \to 0^+$,

$$\begin{aligned}&\langle \eta_t^n, \phi\rangle_Q \to \langle \xi_t, \phi\rangle_Q, \qquad \epsilon^n \langle \triangle \eta^n, \phi\rangle_Q = \epsilon^n \langle \eta^n, \triangle \phi\rangle_Q \to 0,\\ &\langle \mathbf{y}^n \cdot \nabla \eta^n, \phi\rangle_Q = -\langle \mathbf{y}^n \eta^n, \nabla \phi\rangle_Q \to -\langle \mathbf{z}\xi, \nabla \phi\rangle_Q = \langle \mathbf{z} \cdot \nabla \xi, \phi\rangle_Q.\end{aligned}$$

Hence $(\mathbf{z}, \xi)$ satisfies the transport equation

$$\xi_t + \mathbf{z} \cdot \nabla \xi = 0, \quad \text{in } \mathcal{C}_T^*.$$

Since $L^2(Q)\cap H^1(H^{-1})$ is compactly embedded into $C([0,T],H^{-1})$, and $\eta_0^{\epsilon^n}\to\eta_0$ in L^2, the initial condition for ξ is η_0. If we restrict our test function to be zero at time T, i.e. $\{\phi:\phi\in C_T,\phi(T)=0\}$, one can check that

$$\int_0^T\int_\Omega\xi(\phi_t+\mathbf{z}\cdot\nabla\phi)\mathrm{d}x\mathrm{d}t+\int_\Omega\eta_0\phi(0)\mathrm{d}x=0.$$

According to Theorem 4.1 in [2], the weak solution ξ is also a renormalized solution and satisfies $\|\xi(t,\cdot)\|=\|\eta_0\|$. By the property of renormalized solutions (choosing $\beta(s)=(s-m)^2$ in [2]), we have $\|\xi(t,\cdot)-m\|=\|\eta_0-m\|$ for a.e. $t\in[0,T]$. Therefore by (3.35), we have

$$\limsup\|\eta^n-m\|_{L^2(Q)}\le\sqrt{T}\lim\|\eta_0^{\epsilon^n}-m\|=\sqrt{T}\|\eta_0-m\|=\|\xi-m\|_{L^2(Q)}. \tag{3.36}$$

Weak lower semi-continuous of norm implies that

$$\liminf\|\eta^n-m\|_{L^2(Q)}\ge\|\xi-m\|_{L^2(Q)}. \tag{3.37}$$

Combining (3.36) and (3.37) leads to $\|\eta^n-m\|_{L^2(Q)}\to\|\xi-m\|_{L^2(Q)}$ and hence $\eta^n\to\xi$ in $L^2(Q)$. Strong convergence of η^n in $L^2(Q)$ and weak convergence of $\nabla\mathbf{y}^n$ in $L^2(Q)$ imply convergence of $\eta^n\nabla\mathbf{y}^n$ in the distribution sense. For test functions $\mathbf{v}\in\mathcal{V}_T$, we have

$$\begin{aligned}&\langle\mathbf{y}_t^n,\mathbf{v}\rangle_Q\to\langle\mathbf{z}_t,\mathbf{v}\rangle_Q,\\&\langle\operatorname{div}(\eta^n\nabla\mathbf{y}^n),\mathbf{v}\rangle_Q=-\langle\eta^n\nabla\mathbf{y}^n,\nabla\mathbf{v}\rangle_Q\\&\quad\to-\langle\xi\nabla\mathbf{z},\nabla\mathbf{v}\rangle_Q=\langle\operatorname{div}(\xi\nabla\mathbf{z}),\mathbf{v}\rangle_Q.\end{aligned}$$

Hence $(\mathbf{z},\xi)$ satisfies the Stokes equation

$$\mathbf{z}_t-\operatorname{div}(\xi(\nabla\mathbf{z}))=\mathbf{f}\quad\text{in }\mathcal{V}_T^*.$$

Since $L^2(\mathbf{V})\cap H^1(\mathbf{V}^*)$ is compactly embedded into $C([0,T],\mathbf{V}^*)$, we have $\mathbf{z}|_{t=0}=\mathbf{y}_0$. Hence $(\mathbf{z},\xi)$ solves the system (3.1)–(3.5). This implies Theorem 3.2.

Corollary 3.4. *Consider the same assumption as in Theorem 3.2. Then we have*

$$\eta^n\to\eta,\quad\text{in }L^p(Q),\ 2\le p<\infty.$$

Proof. Firstly we observe that if (3.11) is satisfied, then $\eta_0\in[m,M]$ implies that

$$\eta_0^\epsilon\to\eta_0,\quad\text{in }L^p(\Omega),\ 2\le p<\infty.$$

After shifting by a constant function m, we denote $\bar\eta^\epsilon=\eta^\epsilon-m$. Hence $\bar\eta^\epsilon$ satisfies Eq. (3.9) with initial condition $\bar\eta^\epsilon|_{t=0}=\eta_0^\epsilon-m$ and zero Dirichlet boundary condition. It is known that $\bar\eta^\epsilon\in L^\infty(Q)\cap L^2(H^1)$. We can multiply $|\bar\eta^\epsilon|^{p-2}\bar\eta^\epsilon$ on both sides of Eq. (3.9) (where η^ϵ is replaced by $\bar\eta^\epsilon$). Since $\nabla(|\bar\eta^\epsilon|^{p-2}\bar\eta^\epsilon)=(p-1)|\bar\eta^\epsilon|^{p-2}\nabla\bar\eta^\epsilon$, we have

$$\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega|\bar\eta^\epsilon|^p\mathrm{d}x+\epsilon(p-1)\int_\Omega|\bar\eta^\epsilon|^{p-2}|\nabla\bar\eta^\epsilon|^2\mathrm{d}x=0.$$

This implies that $\|\bar\eta^\epsilon(t)\|_{L^p(\Omega)}\le\|\eta_0^\epsilon-m\|_{L^p(\Omega)}$. After passage to the limit and by the same argument as in the proof of Theorem 3.2, we have $\eta^n\rightharpoonup\eta$ in $L^p(Q)$ and η is a renormalized solution. From the property of renormalized solutions (choosing $\beta(s)=|s-m|^p$ in [2]), we have $\|\eta(t,\cdot)-m\|_{L^p(\Omega)}=\|\eta_0-m\|_{L^p(\Omega)}$, a.e. $t\in[0,T]$. Therefore

$$\limsup\|\bar\eta^n-m\|_{L^p(Q)}\le T^{1/p}\lim\|\eta_0^{\epsilon^n}-m\|_{L^p(\Omega)}=T^{1/p}\|\eta_0-m\|_{L^p(\Omega)}=\|\eta-m\|_{L^p(Q)}.$$

Together with weak lower semi-continuous of norm, this implies that $\|\eta^n-m\|_{L^p(Q)}\to\|\eta-m\|_{L^p(Q)}$ and hence $\eta^n\to\eta$ in $L^p(Q)$. □

4. Optimal control formulation

The abstract formulation for the optimal control is:

$$\min J(\mathbf{x},\mathbf{u}),\quad\text{such that }e(\mathbf{x},\mathbf{u})=0, \tag{4.1}$$

where $J(\mathbf{x},\mathbf{u})$ is the cost functional, $\mathbf{u}$ and $\mathbf{x}$ are the control and state variables respectively, and $e(\mathbf{x},\mathbf{u})=0$ denotes the underlying equation. For our control problem, the state variable $\mathbf{x}=(\eta,\mathbf{y})$ and the control variable is $\mathbf{u}$ itself. The underlying equation is either system (3.1)–(3.5) or (3.6)–(3.10). The cost function is

$$J(\eta, \mathbf{u}) = \frac{1}{2}\|\eta - \widetilde{\eta}\|^2_{L^2(Q)} + \frac{\alpha}{2}\|\mathcal{B}\mathbf{u}\|^2_{\mathbf{L}^2(Q)}, \tag{4.2}$$

where $\widetilde{\eta}$ is a given function in $L^2(Q)$. The optimal control problem associated with the original equation is

Problem 4.1.

$\min J(\eta, \mathbf{u})$, such that Eqs. (3.1)–(3.5) hold.

We have existence for the above problem.

Theorem 4.2. *Given $\mathbf{y}_0 \in \mathbf{H}$, $m \le \eta_0 \le M$ a.e., there exists at least one optimal solution $(\eta, \mathbf{y}, \mathbf{u})$ for Problem 4.1, and $(\eta, \mathbf{y}, \mathbf{u}) \in L^\infty(Q) \cap H^1(H^{-1}) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*) \times L^2(\mathbf{L}^2(\tilde{\Omega}))$.*

Proof. Clearly Problem 4.1 is feasible, hence we can find a minimal sequence $(\mathbf{y}^n, \eta^n, \mathbf{u}^n)$, i.e. $\lim_{n\to\infty} J(\eta^n, \mathbf{u}^n) = \inf J(\eta, \mathbf{u})$ such that $(\eta^n, \mathbf{y}^n, \mathbf{u}^n)$ solve (3.1)–(3.5) (we use the equivalent weak form to avoid the pressure term). By the definition of J in (4.2), we know that $\{\mathcal{B}\mathbf{u}^n\}$ is bounded in $\mathbf{L}^2(Q)$. Hence $\{\mathbf{u}^n\}$, $\{\mathbf{y}^n\}$, $\{\eta^n\}$ are also bounded in $L^2(\mathbf{L}^2(\tilde{\Omega}))$, $L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*)$ and $L^\infty(Q) \cap H^1(H^{-1})$ respectively. After passing to the subsequence (still denote the subscript by n), we have

$$\begin{aligned} &\mathbf{u}^n \rightharpoonup \mathbf{u} \quad \text{in } L^2(\mathbf{L}^2(\tilde{\Omega})),\\ &\mathbf{y}^n \rightharpoonup \mathbf{y} \quad \text{in } L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*),\\ &\eta^n \rightharpoonup \eta \quad \text{weak star in } L^\infty(Q) \cap H^1(H^{-1}). \end{aligned}$$

We need to check that $(\mathbf{y}, \eta, \mathbf{u})$ satisfy Eqs. (3.1)–(3.5). Since $\mathbf{y}^n \to \mathbf{y}$ in $\mathbf{L}^2(Q)$ by Aubin's Lemma and $\eta^n \rightharpoonup \eta$ in $L^2(Q)$, we obtain that $\mathbf{y}^n\eta^n$ converges to $\mathbf{y}\eta$ in the distribution sense, hence the transport equation (3.4) is satisfied. The initial condition can be obtained in a similar way as in Theorem 3.2. Hence η is a renormalized solution. By the property of renormalized solution (cf. [2]), $\|\eta^n\|_{L^2(Q)} = \sqrt{T}\|\eta_0\| = \|\eta\|_{L^2(Q)}$. This implies that $\eta^n \to \eta$ in $L^2(Q)$. Together with $\nabla\mathbf{y}^n \rightharpoonup \nabla\mathbf{y}$ in $L^2(Q)$, we have convergence of $\eta^n\nabla\mathbf{y}^n$ in the distribution sense. Therefore the Stokes equations (3.1)–(3.2) are also satisfied. The initial condition can be treated similarly as before. Therefore, $\mathcal{B}\mathbf{u}^n \rightharpoonup \mathcal{B}\mathbf{u}$ in $\mathbf{L}^2(Q)$ implies that $(\eta, \mathbf{y}, \mathbf{u})$ satisfies the underlying equation. Lastly, since the norm $\|\cdot\|_{L^2(Q)}$ is a weakly lower semi-continuous functional, we obtain that $(\eta, \mathbf{u})$ provides a minimum for Problem 4.1. □

Now we move to the optimal control problem associated with the approximated system. We assume that $\mathbf{y}_0 \in \mathbf{V}$, $\eta_0^\epsilon - m \in H_0^1$, then Eqs. (3.6)–(3.10) are equivalent to Eqs. (3.14)–(3.15) with the same initial condition. To avoid the pressure term, we will use Eqs. (3.14)–(3.15). Denote

$$\mathrm{e}^\epsilon(\eta, \mathbf{y}, \mathbf{u}) = \begin{pmatrix} \mathbf{e}^\epsilon_{1,1}, & \mathbf{e}^\epsilon_{1,2} \\ e^\epsilon_{2,1}, & e^\epsilon_{2,2} \end{pmatrix} = \begin{pmatrix} \mathbf{y}_t - P\mathrm{div}\,(\eta\nabla\mathbf{y}) - P\mathcal{B}\mathbf{u}, & \mathbf{y}(0) - \mathbf{y}_0 \\ \eta_t - \epsilon\triangle\eta + \mathbf{y}\cdot\nabla\eta, & \eta(0) - \eta_0^\epsilon \end{pmatrix}. \tag{4.3}$$

Then the optimal control problems for the approximated system are given by:

Problem 4.3.

$\min J(\eta, \mathbf{u})$, such that $\mathrm{e}^\epsilon(\eta, \mathbf{y}, \mathbf{u}) = 0$.

Similar to Theorem 4.2, we have existence for optimal solution for Problem 4.3.

Theorem 4.4. *Given $\mathbf{y}_0 \in \mathbf{V}$, $\eta_0^\epsilon - m \in H_0^1$ and $m \le \eta_0^\epsilon \le M$ a.e., there exists at least one optimal solution $(\eta, \mathbf{y}, \mathbf{u})$ for Problem 4.3, and $(\eta - m, \mathbf{y}, \mathbf{u}) \in L^2(H^2 \cap H_0^1) \cap L^\infty(Q) \cap H^1(L^2) \times L^2(\mathbf{V}) \cap H^1(\mathbf{V}^*) \times L^2(\mathbf{L}^2(\tilde{\Omega}))$.*

In the remainder of this section, we discuss the optimality system for Problem 4.3 and how Problem 4.3 approximates Problem 4.1 when $\epsilon \to 0^+$. Section 4.1 contains the regularity estimates for Eqs. (3.6)–(3.10). The theorems are stated in Section 4.2, i.e. Theorem 4.9 for the optimality system and Proposition 4.10 for the limit property.

4.1. Regularity estimates

We will repeatedly use the following estimates.

Lemma 4.5.

$$\mathbf{y} \in L^2(\mathbf{H}^1) \cap L^\infty(\mathbf{L}^2),\ \eta \in L^2(H^2) \cap L^\infty(H^1) \Rightarrow \mathbf{y}\cdot\nabla\eta \in L^2(Q), \tag{4.4}$$

$$\mathbf{y} \in L^\infty(\mathbf{H}^1),\ \eta \in L^2(H^2) \cap L^\infty(H^1) \Rightarrow \mathbf{y}\cdot\nabla\eta \in L^3(Q), \tag{4.5}$$

$$\mathbf{y} \in L^2(\mathbf{H}^2) \cap L^\infty(\mathbf{H}^1),\ \eta \in L^2(H^2) \cap L^\infty(H^1) \Rightarrow \nabla\eta\cdot\nabla\mathbf{y} \in L^2(Q). \tag{4.6}$$

Proof. The conclusions are based on the Hölder and Sobolev inequalities.

$$\begin{aligned}\int_0^T\int_\Omega |\mathbf{y}\cdot\nabla\eta|^2\mathrm{d}x\mathrm{d}t &= \int_0^T \|\mathbf{y}\cdot\nabla\eta\|_{L^2}^2\mathrm{d}t \le \int_0^T \|\mathbf{y}\|_{L^4}^2\|\nabla\eta\|_{L^4}^2\mathrm{d}t\\ &\le C\int_0^T \|\mathbf{y}\|\,\|\mathbf{y}\|_1\|\eta\|_1\|\eta\|_2\mathrm{d}t \le C(\|\mathbf{y}\|_{L^\infty(\mathbf{L}^2)}^2\|\mathbf{y}\|_{L^2(\mathbf{H}^1)}^2+\|\eta\|_{L^\infty(H^1)}^2\|\eta\|_{L^2(H^2)}^2),\\ \int_0^T\int_\Omega |\mathbf{y}\cdot\nabla\eta|^3\mathrm{d}x\mathrm{d}t &= \int_0^T \|\mathbf{y}\cdot\nabla\eta\|_{L^3}^3\mathrm{d}t \le \int_0^T \|\mathbf{y}\|_{L^6}^3\|\nabla\eta\|_{L^6}^3\mathrm{d}t\\ &\le C\int_0^T \|\mathbf{y}\|_1^3\|\eta\|_1\|\eta\|_2^2\mathrm{d}t,\\ \int_0^T\int_\Omega |\nabla\eta\cdot\nabla\mathbf{y}|^2\mathrm{d}x\mathrm{d}t &= \int_0^T \|\nabla\eta\cdot\nabla\mathbf{y}\|^2\mathrm{d}t \le C\int_0^T \|\nabla\mathbf{y}\|_{L^4}^2\|\nabla\eta\|_{L^4}^2\mathrm{d}t\\ &\le C\int_0^T \|\mathbf{y}\|_1\|\mathbf{y}\|_2\|\eta\|_1\|\eta\|_2\mathrm{d}t \le C\left(\|\mathbf{y}\|_{L^\infty(\mathbf{H}^1)}^2\|\mathbf{y}\|_{L^2(\mathbf{H}^2)}^2+\|\eta\|_{L^\infty(H^1)}^2\|\eta\|_{L^2(H^2)}^2\right).\quad\square\end{aligned}$$

To obtain a higher order regularity result for the solution to the system (3.6)–(3.10), we need to assume that the initial data $\mathbf{y}_0$ and η_0^ϵ satisfying higher regularity properties. Besides (3.11), let

$$\mathbf{y}_0\in\mathbf{V},\qquad \eta_0^\epsilon-m\in H_0^1(\Omega)\cap W^{1,3}(\Omega). \tag{4.7}$$

The existence result (Theorem 3.1) implies that for any $\mathbf{u}\in L^2(\mathbf{H})$, there exists at least one $(\eta,\mathbf{y})$ satisfies (3.6)–(3.10) such that

$$\mathbf{y}\in L^2(\mathbf{V})\cap L^\infty(\mathbf{H})\cap H^1(\mathbf{V}^*),\quad \eta\in L^2(H^1)\cap L^\infty(Q)\cap H^1(H^{-1}). \tag{4.8}$$

Taking the inner product of (3.9) with $-\triangle\eta$, using

$$(\mathbf{y}\cdot\nabla\eta,\triangle\eta)\le\|\mathbf{y}\|_{L^4}\|\nabla\eta\|_{L^4}\|\eta\|_2\le C_\epsilon\|\mathbf{y}\|^2\|\mathbf{y}\|_1^2\|\eta\|_1^2+\frac{1}{2\epsilon}\|\triangle\eta\|^2,$$

and estimate (4.8) and $\eta_0^\epsilon\in H^1$, we have $\eta\in L^2(H^2)\cap L^\infty(H^1)$. Hence $\eta_t\in L^2(Q)$, and

$$\eta\in L^2(H^2)\cap L^\infty(H^1)\cap H^1(L^2). \tag{4.9}$$

After taking the time derivative in (3.9)

$$\eta_{tt}-\epsilon\triangle\eta_t+\mathbf{y}_t\cdot\nabla\eta+\mathbf{y}\cdot\nabla\eta_t=0,$$

and taking the inner product with η_t in the above equation,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\eta_t\|+\epsilon\|\nabla\eta_t\|^2+\langle\mathbf{y}_t\cdot\nabla\eta,\eta_t\rangle=0.$$

Since

$$\langle\mathbf{y}_t\cdot\nabla\eta,\eta_t\rangle=-\langle\mathbf{y}_t\eta,\nabla\eta_t\rangle\le\frac{\epsilon}{2}\|\nabla\eta_t\|^2+C_\epsilon\|\mathbf{y}_t\|^2,$$

we find

$$\|\eta_t\|^2\le C_\epsilon\int_0^t\|\mathbf{y}_t\|^2+\|\eta_t(0)\|^2, \tag{4.10}$$

where $\|\eta_t(0)\|\le\epsilon\|\eta_0^\epsilon\|_2+\|\mathbf{y}_0\cdot\nabla\eta_0^\epsilon\|\le\epsilon\|\eta_0^\epsilon\|_2+C\|\mathbf{y}_0\|_1\|\eta_0^\epsilon\|_2$. Moving η_t in Eq. (3.9) to the right hand side, the elliptic estimation gives

$$\|\triangle\eta\|\le C_\epsilon(\|\eta_t\|+\|\mathbf{y}\|_1). \tag{4.11}$$

Now we move to higher order regularity estimates for the Stokes equation. First we consider the time independent Stokes equation with nonconstant viscosity η.

Lemma 4.6. *Suppose that $\eta\in H^2$ and $m\le\eta\le M$ a.e., and that $\mathbf{y}$ solves following Stokes equation*

$$-\mathrm{div}\,(\eta(\nabla\mathbf{y}))+\nabla p=\mathbf{f}, \tag{4.12}$$

$$\mathrm{div}\,\mathbf{y}=0,\qquad \mathbf{y}|_\Omega=\mathbf{0}. \tag{4.13}$$

Then we have

$$\|\mathbf{y}\|_2\le C(\|\mathbf{f}\|+\|\eta\|_1\|\eta\|_2\|\mathbf{f}\|_{\mathbf{V}^*}). \tag{4.14}$$

Proof. First we have

$$\|\mathbf{y}\|_1 + \|p\| \le C\|\mathbf{f}\|_{\mathbf{V}^*}. \tag{4.15}$$

Eq. (4.12) can be rewritten as

$$-\triangle\mathbf{y} - \frac{\nabla\eta}{\eta}\cdot\nabla\mathbf{y} + \nabla\left(\frac{p}{\eta}\right) + \frac{p\nabla\eta}{\eta^2} = \frac{\mathbf{f}}{\eta}.$$

We multiply $\Lambda\mathbf{y} = -P\triangle\mathbf{y}$ on both sides of the above equation. Since

$$\left\langle\frac{\nabla\eta}{\eta}\cdot\nabla\mathbf{y}, \Lambda\mathbf{y}\right\rangle \le C\|\eta\|_1^{1/2}\|\eta\|_2^{1/2}\|\mathbf{y}\|_1^{1/2}\|\mathbf{y}\|_2^{1/2}\|\Lambda\mathbf{y}\|,$$
$$\left\langle\frac{p\nabla\eta}{\eta^2}, \Lambda\mathbf{y}\right\rangle \le C\|p\|^{1/2}\|p\|_1^{1/2}\|\eta\|_1^{1/2}\|\eta\|_2^{1/2}\|\Lambda\mathbf{y}\|,$$
$$\|p\|_1 \le \|\mathbf{f}\| + M\|\mathbf{y}\|_2 + \|\nabla\eta\|_{L^4}\|\nabla\mathbf{y}\|_{L^4} \le \|\mathbf{f}\| + C\|\mathbf{y}\|_2 + C\|\eta\|_1^{1/2}\|\eta\|_2^{1/2}\|\mathbf{y}\|_1^{1/2}\|\mathbf{y}\|_2^{1/2},$$

together with estimate (4.15) and the property of Stokes operator $\|\mathbf{y}\|_2 \le C\|\Lambda\mathbf{y}\|$, we have

$$\left\langle\frac{\nabla\eta}{\eta}\cdot\nabla\mathbf{y}, \Lambda\mathbf{y}\right\rangle \le \frac{1}{4}\|\Lambda\mathbf{y}\|^2 + C\|\eta\|_1^2\|\eta\|_2^2\|\mathbf{f}\|_{\mathbf{V}^*}^2,$$
$$\left\langle\frac{p\nabla\eta}{\eta^2}, \Lambda\mathbf{y}\right\rangle \le \frac{1}{4}\|\Lambda\mathbf{y}\|^2 + C\|\eta\|_1^2\|\eta\|_2^2\|\mathbf{f}\|_{\mathbf{V}^*}^2,$$
$$\left\langle\frac{\mathbf{u}}{\eta}, \Lambda\mathbf{y}\right\rangle \le \frac{1}{4}\|\Lambda\mathbf{y}\|^2 + C\|\mathbf{f}\|^2.$$

Hence $\|\mathbf{y}\|_2 \le C(\|\mathbf{f}\| + \|\eta\|_1\|\eta\|_2\|\mathbf{f}\|_{\mathbf{V}^*})$. □

We next consider the time dependent Stokes equations (3.6)–(3.8). Multiplying $\mathbf{y}_t$ on both sides of (3.6) and noticing that $(\mathbf{y}_t, \nabla p) = 0$, we have

$$\|\mathbf{y}_t\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\eta\nabla\mathbf{y}, \nabla\mathbf{y}) - \frac{1}{2}\langle\eta_t\nabla\mathbf{y}, \nabla\mathbf{y}\rangle = (\mathcal{B}\mathbf{u}, \mathbf{y}_t).$$

After moving $\mathbf{y}_t$ to the right hand side, using Lemma 4.6 and

$$\|\mathbf{y}_t\|_{\mathbf{V}^*} \le \|\mathcal{B}\mathbf{u}\|_{\mathbf{V}^*} + C\|\mathbf{y}\|_1,$$

we find

$$\|\mathbf{y}\|_2 \le C(\|\mathcal{B}\mathbf{u}\| + \|\mathbf{y}_t\|) + C\|\eta\|_1\|\eta\|_2(\|\mathcal{B}\mathbf{u}\|_{\mathbf{V}^*} + \|\mathbf{y}\|_1).$$

We also have the estimates

$$\begin{aligned}\langle\eta_t\nabla\mathbf{y}, \nabla\mathbf{y}\rangle &\le C\|\eta_t\|\,\|\mathbf{y}\|_1\|\mathbf{y}\|_2 \le C\|\eta_t\|\,\|\mathbf{y}\|_1(\|\mathcal{B}\mathbf{u}\| + \|\mathbf{y}_t\|) + C\|\eta_t\|\,\|\mathbf{y}\|_1\|\eta\|_1\|\eta\|_2(\|\mathcal{B}\mathbf{u}\|_{\mathbf{V}^*} + \|\mathbf{y}\|_1)\\ &\le C\|\eta_t\|^2\|\mathbf{y}\|_1^2 + \|\mathcal{B}\mathbf{u}\|^2 + \frac{1}{2}\|\mathbf{y}_t\|^2 + C(\|\mathbf{y}\|_1^2\|\eta\|_1^2\|\eta\|_2^2 + \|\eta_t\|^2\|\mathcal{B}\mathbf{u}\|^2 + \|\eta_t\|\,\|\eta\|_1\|\eta\|_2\|\mathbf{y}\|_1^2).\end{aligned}$$

Defining $\beta(t) = \int_0^t \|\mathbf{y}_t\|^2 + \|\eta_t(0)\|^2$, $\gamma(t) = (\eta\nabla\mathbf{y}, \nabla\mathbf{y})$, inequality (4.10) implies that $\|\eta_t\| \le C_\epsilon\beta(t)$. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}(\beta+\gamma) \le C_\epsilon(\|\eta_t\|^2 + \|\eta\|_1^2\|\eta\|_2^2\beta + \|\mathcal{B}\mathbf{u}\|^2\gamma) + \|\mathbf{u}\|^2 \le C_\epsilon(\|\eta_t\|^2 + \|\eta\|_1^2\|\eta\|_2^2 + \|\mathcal{B}\mathbf{u}\|^2)(\beta+\gamma) + \|\mathcal{B}\mathbf{u}\|^2.$$

Since we already have (4.9) and $\mathcal{B}\mathbf{u} \in \mathbf{L}^2(Q)$, then by Gronwall inequality, we have $\mathbf{y}_t \in L^2(Q)$ and $\mathbf{y} \in L^\infty(\mathbf{V})$. Hence $\eta_t \in L^\infty(L^2)$ and $\eta \in L^\infty(H^2)$ (from (4.10) and (4.11)), which immediately gives $\mathbf{y} \in L^2(\mathbf{H}^2)$. Summing up we have

$$\mathbf{y} \in L^2(\mathbf{H}^2) \cap L^\infty(\mathbf{V}) \cap H^1(\mathbf{H}). \tag{4.16}$$

To guarantee the existence for a Lagrange multiplier, we need a slight better estimation for η. Due to (4.9) and (4.16), Lemma 4.5 implies $\mathbf{y}\cdot\nabla\eta \in L^3(Q)$. Then moving $\mathbf{y}\cdot\nabla\eta$ to right hand side for Eq. (3.9), and using Theorem 1.14 in [13] and assumption (4.7), we obtain

$$\eta \in L^3(W^{2,3}), \qquad \eta_t \in L^3(Q). \tag{4.17}$$

Combining the results, we find

Theorem 4.7. *Assume that (3.11) and (4.7) are satisfied. If $\mathcal{B}\mathbf{u} \in \mathbf{L}^2(Q)$, then every solution of systems (3.6)–(3.10) satisfies:*

$$\mathbf{y} \in L^2(\mathbf{H}^2) \cap L^\infty(\mathbf{V}) \cap H^1(\mathbf{H}), \qquad \eta \in L^3(W^{2,3}), \qquad \eta_t \in L^3(Q),$$

and

$$\|\mathbf{y}\|_{L^2(\mathbf{H}^2)\cap L^\infty(\mathbf{V})\cap H^1(\mathbf{H})} + \|\eta\|_{L^3(W^{2,3})} + \|\eta_t\|_{L^3(Q)} \le C(\|\mathcal{B}\mathbf{u}\|_{L^2(Q)}),$$

where $C(\cdot)$ maps bounded set to bounded set.

4.2. Optimal control problem associated with approximated equations

This section is devoted to deriving the first order optimality condition for Problem 4.3. Define the spaces

$$X_1 = \{\phi : \phi \in L^3(W^{2,3}),\ \phi_t \in L^3(Q)\}, \tag{4.18}$$

$$\mathbf{Y}_1 = \{\mathbf{v} : \mathbf{v} \in L^2(\mathbf{H}^2) \cap L^\infty(\mathbf{V}) \cap H^1(\mathbf{H})\}. \tag{4.19}$$

By a standard embedding result (cf. [13]), $X_1 \subset L^\infty$. Recalling (4.3) for the definition of the nonlinear map e^ϵ, we have

Lemma 4.8. *For any fixed positive constant ϵ, the map e^ϵ acts from $X_1 \times \mathbf{Y}_1 \times L^2(\mathbf{L}^2(\tilde{\Omega}))$ to $\begin{pmatrix} L^2(\mathbf{H}), & \mathbf{V} \\ L^3(Q), & W^{1,3} \end{pmatrix}$. Moreover, it is Frechet differentiable.*

Proof. We first verify that e is well-defined. Recalling that $\gamma_0(\eta) = \eta(0)$ and $\gamma_0(\mathbf{y}) = \mathbf{y}(0)$. Here γ_0 is continuous from X_1 to $W^{1,3}$ (see e.g. Theorem 1.13 in [13]) and also continuous from $\mathbf{Y}_2$ to $\mathbf{V}$ (see e.g. [8]), then $\mathrm{e}^\epsilon_{1,2}$ and $\mathrm{e}^\epsilon_{2,2}$ are well defined. For any given $(\eta, \mathbf{y}, \mathbf{u}) \in X_1 \times \mathbf{Y}_1 \times L^2(\mathbf{L}^2(\tilde{\Omega}))$, by virtue of $X_1 \hookrightarrow L^\infty(Q)$ and Lemma 4.5, we have

$$(\mathrm{div}\ (\eta \nabla \mathbf{y})) = \eta \triangle \mathbf{y} + \nabla \eta \cdot \nabla \mathbf{y} \in \mathbf{L}^2(Q), \quad \mathbf{y} \cdot \nabla \eta \in L^3(Q).$$

Hence $\mathrm{e}^\epsilon_{1,1}$ and $\mathrm{e}^\epsilon_{2,1}$ lie in $L^2(\mathbf{H})$ and $L^3(Q)$ respectively. Since $\mathrm{e}^\epsilon_{1,2}$ and $\mathrm{e}^\epsilon_{2,2}$ are linear operators, the differentiability is clear. For $\mathrm{e}^\epsilon_{1,1}$ and $\mathrm{e}^\epsilon_{2,1}$, consider the linearized equation at point $(\eta, \mathbf{y}, \mathbf{u})$ as

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \begin{pmatrix} \mathrm{e}^\epsilon_{1,1}(\eta, \mathbf{y}, \mathbf{u}) \\ \mathrm{e}^\epsilon_{2,1}(\eta, \mathbf{y}, \mathbf{u}) \end{pmatrix} (\delta\eta, \delta\mathbf{y}, \delta\mathbf{u}) = \begin{pmatrix} \delta\mathbf{y}_t - P(\mathrm{div}\ (\delta\eta \nabla \mathbf{y})) - P(\mathrm{div}\ (\eta \nabla \delta\mathbf{y})) - P\mathcal{B}\delta\mathbf{u} \\ \delta\eta_t - \epsilon \triangle \delta\eta + \delta\mathbf{y} \cdot \nabla \eta + \mathbf{y} \cdot \nabla \delta\eta \end{pmatrix}. \tag{4.20}$$

We will check that the linearized equation is indeed the Frechet derivative. By calculation,

$$\begin{pmatrix} \mathrm{e}^\epsilon_{1,1}(\eta + \delta\eta, \mathbf{y} + \delta\mathbf{y}, \mathbf{u} + \delta\mathbf{u}) \\ \mathrm{e}^\epsilon_{2,1}(\eta + \delta\eta, \mathbf{y} + \delta\mathbf{y}, \mathbf{u} + \delta\mathbf{u}) \end{pmatrix} - \begin{pmatrix} \mathrm{e}^\epsilon_{1,1}(\eta, \mathbf{y}, \mathbf{u}) \\ \mathrm{e}^\epsilon_{2,1}(\eta, \mathbf{y}, \mathbf{u}) \end{pmatrix} - \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \begin{pmatrix} \mathrm{e}^\epsilon_{1,1}(\eta, \mathbf{y}, \mathbf{u}) \\ \mathrm{e}^\epsilon_{2,1}(\eta, \mathbf{y}, \mathbf{u}) \end{pmatrix} (\delta\eta, \delta\mathbf{y}, \delta\mathbf{u}) = \begin{pmatrix} P(\mathrm{div}\ (\delta\eta \nabla \delta\mathbf{y})) \\ \delta\mathbf{y} \cdot \nabla \delta\eta \end{pmatrix}.$$

Since

$$\|P(\mathrm{div}\ (\delta\eta \nabla \delta\mathbf{y}))\|_{L^2(Q)} \le \|\delta\eta\|_{L^\infty(Q)} \|\triangle \delta\mathbf{y}\|_{L^2(Q)} + \|\nabla \delta\eta \cdot \nabla \delta\mathbf{y}\|_{\mathbf{L}^2(Q)},$$

by $X_1 \hookrightarrow L^\infty(Q)$ and Lemma 4.5, we have

$$\|P(\mathrm{div}\ (\delta\eta \nabla \delta\mathbf{y}))\|_{L^2(Q)} + \|\delta\mathbf{y} \cdot \nabla \delta\eta\|_{L^3(Q)} \le C \|\delta\eta\|_{X_1} \|\delta\mathbf{y}\|_{\mathbf{Y}_1}.$$

Recalling the definition of the Frechet derivative, we conclude that e^ϵ is differentiable with derivative e^ϵ_x. □

The existence of an optimal solution for Problem 4.3 was already obtained in Theorem 4.4. We let $(\eta^*_\epsilon, \mathbf{y}^*_\epsilon, \mathbf{u}^*_\epsilon)$ be one optimal solution. From Lemma 4.8, e^ϵ is differentiable, and hence $\mathrm{e}^\epsilon_x(\eta^*_\epsilon, \mathbf{y}^*_\epsilon, \mathbf{u}^*_\epsilon)$ also maps $X_1 \times \mathbf{Y}_1 \times L^2(\mathbf{L}^2(\tilde{\Omega}))$ to $\begin{pmatrix} L^2(\mathbf{H}), & \mathbf{V} \\ L^3(Q), & W^{1,3} \end{pmatrix}$. Moreover, this map is also surjective. In fact, for any $\begin{pmatrix} \mathbf{g}_1, & \mathbf{q}_1 \\ g_2, & q_2 \end{pmatrix} \in \begin{pmatrix} L^2(\mathbf{H}), & \mathbf{V} \\ L^3(Q), & W^{1,3} \end{pmatrix}$, we verify that there exists $(\delta\eta, \delta\mathbf{y}, \delta\mathbf{u})$ which satisfies:

$$\begin{pmatrix} \delta\mathbf{y}_t - P(\mathrm{div}\ (\delta\eta \nabla \mathbf{y}^*_\epsilon)) - P(\mathrm{div}\ (\eta^*_\epsilon \nabla \delta\mathbf{y})) - P\mathcal{B}\delta\mathbf{u} = \mathbf{g}_1 \\ \delta\eta_t - \epsilon \triangle \delta\eta + \delta\mathbf{y} \cdot \nabla \eta^*_\epsilon + \mathbf{y}^*_\epsilon \cdot \nabla \delta\eta = g_2 \end{pmatrix}, \tag{4.21}$$

with initial condition

$$\begin{pmatrix} \delta\mathbf{y}(0) = \mathbf{q}_1 \\ \delta\eta(0) = q_2 \end{pmatrix}. \tag{4.22}$$

One can choose $\delta\mathbf{u} = 0$, then (4.21)–(4.22) is a coupled linear parabolic system. By a similar argument as in the previous subsection, we have $\delta\mathbf{y} \in \mathbf{Y}_1$ and $\delta\eta \in L^2(H^2) \cap L^\infty(H^1)$. By Lemma 4.5, $\delta\mathbf{y} \cdot \nabla \eta^*_\epsilon \in L^3(Q)$. Let $\delta\eta$ solve the equation

$$\delta\eta_t - \epsilon \triangle \delta\eta + \mathbf{y}^*_\epsilon \cdot \nabla \delta\eta = g_2 - \delta\mathbf{y} \cdot \nabla \eta^*_\epsilon$$

with initial condition q_2 and zero boundary condition. A similar argument as in Theorem 4.7 implies $\delta\eta \in X_1$ and surjectivity follows. The surjectivity of $e_x^\epsilon(\eta_\epsilon^*, \mathbf{y}_\epsilon^*, \mathbf{u}_\epsilon^*)$ implies there exists a Lagrange multiplier $(\mathbf{z}, \xi) \in L^2(\mathbf{H}) \times L^{4/3}(Q)$, such that the following Lagrangian

$$\mathcal{L}(\eta, \mathbf{y}, \mathbf{u}, \xi, \mathbf{z}) = \frac{1}{2}\|\eta - \widetilde{\eta}\|^2_{L^2(Q)} + \frac{\alpha}{2}\|\mathcal{B}\mathbf{u}\|^2_{L^2(Q)} + (\mathbf{z}, \mathbf{y}_t - P(\mathrm{div}\,(\eta\nabla\mathbf{y})) - P\mathcal{B}\mathbf{u}) + \langle \xi, \eta_t - \epsilon\triangle\eta + \mathbf{y}\cdot\nabla\eta\rangle_{L^{4/3}(Q),L^3(Q)} \tag{4.23}$$

has a stationary point $(\eta_\epsilon^*, \mathbf{y}_\epsilon^*, \mathbf{u}_\epsilon^*, \xi, \mathbf{z})$, see e.g. [14]. In particular, we have

$$\mathcal{L}_\xi = 0,\ \mathcal{L}_\mathbf{z} = 0 \Rightarrow \text{primal equation},$$
$$\mathcal{L}_\eta = 0,\ \mathcal{L}_\mathbf{y} = 0 \Rightarrow \text{adjoint equation},$$
$$\mathcal{L}_\mathbf{u} = 0 \Rightarrow \text{optimal condition.}$$

Expressing these facts in PDE form we obtain the following result.

Theorem 4.9. *Let $(\eta_\epsilon^*, \mathbf{y}_\epsilon^*, \mathbf{u}_\epsilon^*)$ denote an optimal solution of Problem 4.3 and let $(\xi, \mathbf{z})$ be an associated Lagrange multiplier. Then they satisfy the following equations.*

Primal Equation:

$$\mathbf{y}_t - P(\mathrm{div}\,(\eta\nabla\mathbf{y})) = P\mathcal{B}\mathbf{u},$$
$$\mathbf{y}|_{\partial\Omega} = \mathbf{0}, \qquad \mathbf{y}|_{t=0} = \mathbf{y}_0,$$
$$\eta_t - \epsilon\triangle\eta + \mathbf{y}\cdot\nabla\eta = 0,$$
$$\eta|_{t=0} = \eta_0^\epsilon, \qquad \eta|_{\partial\Omega} = m.$$

Adjoint Equation (in the weak sense):

$$-\mathbf{z}_t - \mathrm{div}\,(\eta\nabla\mathbf{z}) - \eta\nabla\xi = \mathbf{0},$$
$$-\xi_t - \epsilon\triangle\xi + \mathbf{y}\cdot\nabla\xi + \nabla\mathbf{y} : \nabla\mathbf{z} = \eta - \widetilde{\eta},$$
$$\mathbf{z}|_{t=T} = \mathbf{0}, \qquad \eta|_{t=T} = 0,$$

where $\nabla\mathbf{y} : \nabla\mathbf{z}$ is the matrix inner product of Frobenius type.

Optimality Condition:

$$\alpha\mathcal{B}^*\mathcal{B}\mathbf{u} = \mathcal{B}^*\mathbf{z},$$

where the projection operator can be ignored due to $\mathbf{z}$ being also divergence free.

Next we consider the relation between the minimum of Problems 4.1 and 4.3. Let $(\eta^*, \mathbf{y}^*, \mathbf{u}^*)$ and $(\eta_\epsilon^*, \mathbf{y}_\epsilon^*, \mathbf{u}_\epsilon^*)$ be a minimizer for Problems 4.1 and 4.3 respectively, and define

$$J^* = J(\eta^*, \mathbf{u}^*), \qquad J^\epsilon = J(\eta_\epsilon^*, \mathbf{u}_\epsilon^*).$$

Then we have

Proposition 4.10.

$$J^* \leq \varliminf_{\epsilon\to 0^+} J^\epsilon.$$

Proof. Consider any convergent subsequence $(\eta_n^*, \mathbf{y}_n^*, \mathbf{u}_n^*)$ for $\epsilon^n \to 0^+$ (we use n to replace ϵ^n for simplicity), and suppose it converges to triple $(\eta, \mathbf{y}, \mathbf{u})$ weak star in the space $L^\infty(Q)\cap H^1(H^{-1}) \times L^2(\mathbf{V})\cap H^1(\mathbf{V}^*) \times L^2(\mathbf{L}^2(\tilde{\Omega}))$. By a similar argument as in Theorem 3.2, one can find that $(\eta, \mathbf{y}, \mathbf{u})$ satisfies (3.1)–(3.5). By definition, we have

$$J^* \leq J(\eta, \mathbf{u}) \leq \varliminf_{\epsilon^n\to 0^+} J^n.$$

Since the above inequality holds for any convergence subsequence with $\epsilon^n \to 0^+$, we have the conclusion. □

The following results involve the set of all solutions to $(\eta, \mathbf{y}, \mathbf{u})$ to (3.1)–(3.5) which are approximated by solutions of the regularized system (3.6)–(3.10).

$$\mathcal{S} = \{(\eta, \mathbf{y}, \mathbf{u}) \in L^\infty(Q)\cap H^1(H^{-1}) \times L^2(\mathbf{V})\cap H^1(\mathbf{V}^*) \times L^2(\mathbf{L}^2(\tilde{\Omega})),\ \text{solution to (3.1)–(3.5)}:$$
$$(\eta, \mathbf{u}) = \lim_{\epsilon_n}(\eta_{\epsilon_n}, \mathbf{u}_{\epsilon_n}) \text{ in } L^2(Q)\times L^2(\mathbf{L}^2(\tilde{\Omega})),\ \text{with } \lim_{n\to\infty}\epsilon_n = 0,\ \text{and } (\eta_\epsilon, y_\epsilon, u) \text{ solution to (3.6)–(3.10)}\}.$$

If for $u \in L^2(\mathbf{L}^2(\tilde{\Omega}))$ the solution $(\eta, \mathbf{y}, \mathbf{u})$ to (3.1)–(3.5) is unique then it is an element of $\mathcal{S}$.

Proposition 4.11. *If there exists an optimal solution* $(\eta^*, \mathbf{y}^*, \mathbf{u}^*)$ *which lies in* $\mathcal{S}$ *then*

$$J^* = \lim_{\epsilon \to 0^+} J^\epsilon.$$

Proof. We have the lower bound from Proposition 4.10. To verify that equality prove that $J^* \geq \overline{\lim}_{\epsilon\to 0^+} J^\epsilon$. Since $(\eta^*, \mathbf{y}^*, \mathbf{u}^*)$ there exists a sequence ϵ^n such that $(\eta^n, \mathbf{y}^n, \mathbf{u}^n)$ solves system (3.6)–(3.10) with $\epsilon = \epsilon^n$, and $(\eta^n, \mathbf{u}^n) \to (\eta^*, \mathbf{u}^*)$ in the strong topology of $L^2(Q) \times L^2(\mathbf{L}^2(\tilde{\Omega}))$. Hence

$$J^n \leq J(\eta^n, \mathbf{u}^n) = \frac{1}{2}\|\eta^n - \widetilde{\eta}\|^2_{L^2(Q)} + \frac{\alpha}{2}\|\mathcal{B}\mathbf{u}^n\|^2_{\mathbf{L}^2(Q)} \to J(\eta^*, \mathbf{u}^*) = J^*$$

By Proposition 4.10 and the notation $u^n = u^{\epsilon_n}$, liminf and limsup concide, and the limit is J^*. □

Acknowledgements

The authors deeply appreciate several useful comments by an anonymous referee. The work of K.K. was in part supported through SFB Mathematical Optimization and Applications in Biomedical Sciences, Fonds zur Förderung der wissenschaftlichen Forschung, Austria.

References

[1] R.J. DiPerna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989) 511–547.
[2] A. Nouri, F. Poupaud, Y. Demay, An existence theorem for the multi-fluid Stokes problem, Quart. Appl. Math. (1997) 421–435.
[3] A. Nouri, F. Poupaud, An existence theorem for the multi-fluid Navier–Stokes problem, J. Differential Equations 122 (1995) 71–88.
[4] H. Abels, On generalized solutions of two-phase flows for viscous incompressible fluids, Interfaces Free Bound. 9 (1) (2007) 31–65.
[5] Y. Giga, S. Takahashi, On global weak solutions of the nonstationary two-phase Stokes flow, SIAM J. Math. Anal. 26 (1994) 876–893.
[6] B. Desjardins, Regularity results for two-dimensional flows of multiphase viscous fluids, Arch. Ration. Mech. Anal. 137 (1997) 135–158.
[7] V. Barbu, Analysis and Control of Nonlinear Infinite Dimensional Systems, Academic Press, Boston, 1993.
[8] R. Temam, Navier–Stokes Equations, North-Holland, 1979.
[9] O.A. Ladyženskaja, V.A. Solonnikov, N.N. Uralćeva, Linear and Quasilinear Equations of Parabolic Type, in: Translations of Mathematical Monographs, AMS, 1968.
[10] M.S. Berger, Nonlinearity and Functional Analysis, Academic Press, 1977.
[11] L.C. Evans, Partial Differential Equations, American Mathematical Society, 1998.
[12] W. Rudin, Principles of Mathematical Analysis, McGraw Hill, 1976.
[13] A.V. Fursikov, Optimal Control of Distributed Systems. Theory and Applications, American Mathematical Society, 1999.
[14] H. Maurer, J. Zowe, First and second order necessary and sufficient optimality conditions for infinited dimensional programming problems, Math. Program. 16 (1979) 98–110.