DIRICHLET CONTROL OF UNSTEADY NAVIER–STOKES TYPE SYSTEM RELATED TO SORET CONVECTION BY BOUNDARY PENALTY METHOD

S.S. Ravindran[1]

Abstract. In this paper, we study the boundary penalty method for optimal control of unsteady Navier–Stokes type system that has been proposed as an alternative for Dirichlet boundary control. Existence and uniqueness of solutions are demonstrated and existence of optimal control for a class of optimal control problems is established. The asymptotic behavior of solution, with respect to the penalty parameter ϵ, is studied. In particular, we prove convergence of solutions of penalized control problem to the corresponding solutions of the Dirichlet control problem, as the penalty parameter goes to zero. We also derive an optimality system and determine optimal solutions. In order to illustrate the theoretical results and the practical utility of control, we numerically address the problem of controlling unsteady convection with Soret effect using a gradient-based method. Numerical results show the effectiveness of the approach.

1. Introduction

Control of fluid flows modeled by Navier–Stokes systems has received considerable attention recently due to its importance in practice and to the theoretical and computational challenges it poses. There is now an extensive body of literature devoted to this subject, see [10, 13, 30] for surveys in this area. In this paper, we consider optimal Dirichlet boundary control of unsteady Navier–Stokes type PDEs. 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, [8, 9, 14, 15, 17]. In [14, 15] control space is taken to be H^1 (smooth controls) leading to a boundary Laplace equation for the control. In [8], a separation of variable type Dirichlet control of unsteady Navier–Stokes equation is studied. In [9], 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 [17], a boundary penalty approach is

[1] Department of Mathematical Sciences, 201C Shelby Center for Science and Technology, The University of Alabama in Huntsville, Huntsville, AL 35899, USA. `ravinds@uah.edu`

studied for stationary Navier–Stokes equations. Starting with the work of [17], interest in penalization technique for treating Dirichlet control has increased. It has since been studied in the context of optimal control with other governing equations [3, 4, 7].

In this paper, we present a boundary penalty approach to treat the Dirichlet control of unsteady Navier–Stokes type system modeling doubly diffusive convection with Soret effect. Significant work has been devoted to studying the stability and physics of doubly diffusive convection with and without Soret effect (thermal diffusion), see for *e.g.* [2, 5, 19, 22, 26, 29]. These studies have reported convective flows lead to undesirable effects in certain applications. For example, thermosolutal convection is responsible for macrosegragation and can affect the uniformity and speed of growth rate in crystal growth. It is also responsible for erosion of gradient zone in solar ponds and roll-over instability (sudden over pressure) in storage and transport of gases. In spite of this, work concerning control of doubly diffusive flows is quite limited although there exists substantial work on control of thermal convection in fluid flows [6, 20, 28], for example. In [32], control of temperature in doubly diffusive flows is studied computationally using boundary heat flux ignoring Soret effect. In this paper, we study control of convection using boundary velocity in doubly diffusive flows with Soret effect. Boundary penalty approach studied here allows one to work with L^2- control space. Unlike other L^2- control space approaches it does not lead to optimality conditions that involve normal derivative of the adjoint variable on the boundary.

The remainder of the paper is organized as follows. In Section 2, we present some mathematical preliminaries and study the wellposedness of the governing equations. In Section 3, we formulate the penalized boundary control problem. The convergence of its solutions to the solutions of the Dirichlet control problem is studied. In Section 4, we derive the optimality system for the optimal control problem. In Section 5, we present results from the numerical implementation. Finally, we list some important estimates in Appendix A.

2. Preliminaries

2.1. Notations

In our analysis, we will employ the following notations and function spaces. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with 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 $H_0^1(\Omega)$. The closed subspace of functions in $L^2(\Omega)$ with zero mean on Ω will be denoted by $L_0^2(\Omega)$. The dual space of $H_0^1(\Omega)$ 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

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

and

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

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 a Banach space X, we denote by $L^p(0, T; X)$ the time-space function space endowed with the norm $\|w\|_{L^p(0,T;X)} := \left(\int_0^T \|w\|_X^p \, dt \right)^{1/p}$ if $1 \leq p < \infty$ and $\text{ess sup}_{t \in [0,T]} \|w\|_X$ 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 $\mathcal{W}(0, T) := \mathbf{W}_1 \times W_2 \times W_2$, where

$$\mathbf{W}_1 := \{ \mathbf{u} \in L^2(\mathbf{V}) : \mathbf{u}_t \in L^2(\mathbb{V}^*) \} \quad \text{and} \quad W_2 := \{ \phi \in L^2(H_0^1(\Omega)) : \phi_t \in L^2(H^{-1}(\Omega)) \}$$

endowed with the norms $\|\mathbf{u}\|_{W_1} := (\|\mathbf{u}\|_{L^2(\mathbf{V})}^2 + \|\mathbf{u}_t\|_{L^2(\mathbb{V}^*)}^2)^{\frac{1}{2}}$ and $\|\phi\|_{W_2} := (\|\phi\|_{L^2(H_0^1)}^2 + \|\phi_t\|_{L^2(H^{-1})}^2)^{\frac{1}{2}}$, respectively. Since $\mathbf{W}_1$ is continuously embedded in $C(0,T;\mathbf{H})$, see [9], we have that $W(0,T)$ is continuously embedded in $C(0,T;\mathbf{H} \times (L^2(\Omega))^2)$. We further define

$$\mathbf{W}_1^{\Sigma} := \{\mathbf{g}|\tau(\mathbf{u}) = \mathbf{g} \text{ for } \mathbf{u} \in \mathbf{W}_1\},$$

where $\tau : \mathbf{W}_1 \to L^2(0,T;\mathbf{H}^{\frac{1}{2}}(\Gamma))$ 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 $\mathbf{W}_1^{\Sigma}$ by

$$\|\mathbf{g}\|_{W_1^{\Sigma}} := \inf_{\substack{\mathbf{u} \in W_1 \\ \tau(\mathbf{u})=\mathbf{g}}} \|\mathbf{u}\|_{W_1}.$$

For $\mathbf{g} \in \mathbf{W}_1^{\Sigma}$, we define by $\mathbf{u}_g$ the unique element in $\mathbf{W}_1$ that achieves this infimum, see [16].

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

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

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

where $\lambda > 0$ is a constant, see ([23], Thm. 1.9) and ([18], Sect. 5.3).

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

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

see, [11, 24]. Notice that with $q = 2$ and $r = 4$, Gagliardo−Nirenberg inequality implies

$$L^2(0,T;X) \cap L^\infty(0,T;Y) \subset L^4(\Omega \times (0,T)), \tag{2.1}$$

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

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

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

2.2. Governing equations and weak formulation

In this section, we present the governing equations and study their well-posedness. The equations for the doubly diffusive convection with Soret effect in a binary mixture may be written, employing a Boussinesq approximation in the body force term in the momentum equation, as

$$\begin{cases} \partial_t\mathbf{u} - Pr\Delta\mathbf{u} + \mathbf{u} \cdot \nabla\mathbf{u} + \nabla p = Pr^2 Gr_\theta(\theta + N_r S)\mathbf{i}_3, \\ \nabla \cdot \mathbf{u} = 0, \\ \partial_t\theta + \mathbf{u} \cdot \nabla\theta - \Delta\theta = f_1, \\ \partial_t S + \mathbf{u} \cdot \nabla S - \dfrac{1}{Le}\Delta S = \dfrac{\alpha_*}{Le}\Delta\theta + f_2. \end{cases} \tag{2.2}$$

with the boundary conditions

$$\mathbf{u}|_\Gamma = \mathbf{g}, \quad \theta|_\Gamma = 0, \quad S|_\Gamma = 0 \tag{2.3}$$

and initial conditions

$$\mathbf{u}(\mathbf{x},0) = \mathbf{u}_0(\mathbf{x}), \qquad \theta(\mathbf{x},0) = \theta_0(\mathbf{x}), \qquad S(\mathbf{x},0) = S_0(\mathbf{x}) \quad \text{in } \Omega, \tag{2.4}$$

where $\mathbf{u}$ is the velocity, θ the temperature, S the concentration and p the pressure. The non-dimensional parameters Pr, Le, Gr_θ and Gr_S denote the Prandtl number, the Lewis number, the thermal Grashof number and the species Grashof number, respectively. The ratio of species buoyancy to thermal buoyancy N_r is defined by $N_r = \frac{Gr_S}{Gr_\theta}$. In $(2.2)_4$ the first term on the right side corresponds to Soret effect. The cases $\alpha_* > 0$ and $\alpha_* < 0$ corresponds to positive and negative Soret effect, respectively. The Soret effects can have significant implications on convection in liquid mixtures, for example semi-conductor crystal growth [19]. Therefore the Dufour effect has been neglected here in comparison to Soret effect as is common for flows in liquid mixture.

Throughout this paper, we make the following assumption about the data and spatial domain.

Assumption (A). *We assume throughout that the domain $\Omega \subset \mathbb{R}^2$ is bounded with Lipschitzian boundary Γ. Moreover, the data satisfy $\mathbf{u}_0 \in \mathbf{H}$, $\theta_0, S_0 \in L^2(\Omega)$, $f_1, f_2 \in L^2(0, T; H^{-1}(\Omega))$.*

We define the weak solutions of the doubly diffusive convection model with Dirichlet boundary conditions as follows.

Definition 2.1. Given $\mathbf{g} \in \mathbf{W}_1^\Sigma$, a triple $(\mathbf{u}, \theta, S) \in \mathcal{W}(0, T)$ is said to be a weak solution of (2.2), (2.3) and (2.4) if

$$
\begin{cases}
(\partial_t \mathbf{u}, \mathbf{v}) + Pr(\nabla \mathbf{u}, \nabla \mathbf{v}) + (\mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v}) = (Pr^2 Gr_\theta (\theta + N_r S) \mathbf{i}_3, \mathbf{v}), \\
\quad (\partial_t \theta, \phi) + (\nabla \theta, \nabla \phi) + (\mathbf{u} \cdot \nabla \theta, \phi) = (f_1, \phi), \\
(\partial_t S, \psi) + \dfrac{1}{Le}(\nabla S, \nabla \psi) + (\mathbf{u} \cdot \nabla S, \psi) = \dfrac{\alpha_*}{Le}(\nabla \theta, \nabla \psi) + (f_2, \psi),
\end{cases}
\tag{2.5}
$$

$$
\mathbf{u}\,|_\Gamma = \mathbf{g}, \ \ \theta\,|_\Gamma = 0, \ \ S\,|_\Gamma = 0
$$

and

$$
\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \theta(\mathbf{x}, 0) = \theta_0(\mathbf{x}) \quad \text{and} \quad S(\mathbf{x}, 0) = S_0(\mathbf{x}),
$$

$\forall (\mathbf{v}, \phi, \psi) \in \mathbb{V} \times (H_0^1(\Omega))^2$, for a.e. $t \in (0, T]$.

Before we define the penalized boundary value problem, we observe that by integration by parts and $\mathbf{H}^{1/2}(\Gamma) \hookrightarrow \mathbf{L}^3(\Gamma)$ [1], the following identity holds for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega)$ with $\nabla \cdot \mathbf{u} = 0$:

$$
((\mathbf{u} \cdot \nabla)\mathbf{v}, \mathbf{w}) = -(\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v}) + (\mathbf{v}, \mathbf{w}(\mathbf{u} \cdot \mathbf{n}))_\Gamma.
$$

If we define a tri-linear form $c(\cdot, \cdot, \cdot)$ 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 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega)$, it is clear that $c(\mathbf{u}, \mathbf{v}, \mathbf{v}) = 0$. Moreover, for $\mathbf{u}, \mathbf{v} \in \mathbf{H}^1(\Omega)$, $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$ with $\nabla \cdot \mathbf{u} = 0$,

$$
c(\mathbf{u}, \mathbf{v}, \mathbf{w}) = ((\mathbf{u} \cdot \nabla)\mathbf{w}, \mathbf{v}) = -((\mathbf{u} \cdot \nabla)\mathbf{v}, \mathbf{w}).
$$

Motivated by the above discussion we define the penalized boundary conditions for the velocity $\mathbf{u}$ as

$$
-p\mathbf{n} + Pr\frac{\partial \mathbf{u}}{\partial \mathbf{n}} - \frac{1}{2}(\mathbf{u} \cdot \mathbf{n})\mathbf{u} + \frac{1}{\epsilon}\mathbf{u} = \frac{1}{\epsilon}\mathbf{g} \quad \text{on } \Gamma.
\tag{2.6}
$$

We now define the the weak solution to the penalized boundary value problem (2.2), (2.4) and (2.6) as follows:

Definition 2.2. Given $\mathbf{g} \in L^2(\mathbf{L}^2(\Gamma))$, a triple $(\mathbf{u}, \theta, S) \in \mathcal{W}(0, T)$ is said to be a weak solution of (2.2), (2.4) and (2.6) if

$$
\begin{cases}
(\partial_t \mathbf{u}, \mathbf{v}) + Pr(\nabla \mathbf{u}, \nabla \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + \frac{1}{\epsilon}(\mathbf{u}, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\theta + N_r S)\mathbf{i}_3, \mathbf{v}) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \frac{1}{\epsilon}(\mathbf{g}, \mathbf{v})_\Gamma, \\
(\partial_t \theta, \phi) + (\nabla \theta, \nabla \phi) + (\mathbf{u} \cdot \nabla \theta, \phi) = (f_1, \phi), \\
(\partial_t S, \psi) + \frac{1}{Le}(\nabla S, \nabla \psi) + (\mathbf{u} \cdot \nabla S, \psi) = \frac{\alpha_*}{Le}(\nabla \theta, \nabla \psi) + (f_2, \psi),
\end{cases}
\tag{2.7}
$$

and

$$
\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \theta(\mathbf{x}, 0) = \theta_0(\mathbf{x}) \quad \text{and} \quad S(\mathbf{x}, 0) = S_0(\mathbf{x}),
$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$.

Lemma 2.3. *Assume $\epsilon \in (0, \frac{1}{Pr})$, $\mathbf{g} \in L^2(0, T; L^2(\Gamma))$. Then, there exists a unique solution $(\mathbf{u}, \theta, S) \in \mathcal{W}(0, T)$ satisfying (2.7) and*

$$
\sup_{t \in [0,T]} \|\theta\|^2 + \|\theta\|^2_{L^2(H_0^1(\Omega))} \leq M_1, \quad \sup_{t \in [0,T]} \|S\|^2 + \frac{1}{Le}\|S\|^2_{L^2(H_0^1(\Omega))} \leq M_2,
$$

and

$$
\sup_{t \in [0,T]} \|\mathbf{u}\|^2 + Pr\|\nabla \mathbf{u}\|^2_{L^2(\mathbf{H})} + \frac{1}{2\epsilon}\|\mathbf{u}\|^2_{L^2(L^2(\Gamma))} \leq M_3 + \frac{2}{\epsilon}\|\mathbf{g}\|^2_{L^2(L^2(\Gamma))}
$$

where $M_1 := \|\theta_0\|^2 + \|f_1\|^2_{L^2(H^{-1}(\Omega))}$, $M_2 := \|S_0\|^2 + 2Le\|f_2\|^2_{L^2(H^{-1})} + \frac{2\alpha_*^2}{Le}M_1$ and $M_3 := \frac{Pr^3 Gr^2}{\lambda}[M_1 + N_r^2 M_2] + \|\mathbf{u}_0\|^2$.

Proof. We employ standard Galerkin method to prove the existence of solutions. Let $\{(\mathbf{e}_k(\mathbf{x}), a_k(\mathbf{x}))\}_{k=1}^\infty$ be an orthogonal basis of $\mathbf{V} \times H_0^1(\Omega)$ such that $\{(\mathbf{e}_k(\mathbf{x}), a_k(\mathbf{x}))\}_{k=1}^\infty$ is orthonormal in $\mathbf{L}^2(\Omega) \times L^2(\Omega)$. For each $m = 1, 2, \ldots$, we set $\mathbf{V}^m := \text{span}\{\mathbf{e}_i\}_{i=1}^m \times (\text{span}\{a_i\}_{i=1}^m)^2$ and let $\mathbf{u}_m = \sum_{k=1}^m c_k^{(m)} \mathbf{e}_k$, $\theta_m = \sum_{k=1}^m d_k^{(m)} a_k$ and $S_m = \sum_{k=1}^m r_k^{(m)} a_k$ be a solution of

$$
\begin{cases}
(\partial_t \mathbf{u}_m, \mathbf{e}_k) + Pr(\nabla \mathbf{u}_m, \nabla \mathbf{e}_k) + c(\mathbf{u}_m, \mathbf{u}_m, \mathbf{e}_k) + \frac{1}{\epsilon}(\mathbf{u}_m, \mathbf{e}_k)_\Gamma \\
\qquad\qquad\qquad\qquad = (Pr^2 Gr_\theta(\theta_m + N_r S_m)\mathbf{i}_3, \mathbf{e}_k) + \frac{1}{\epsilon}(\mathbf{g}, \mathbf{e}_k)_\Gamma, \\
(\partial_t \theta_m, a_k) + (\nabla \theta_m, \nabla a_k) + (\mathbf{u}_m \cdot \nabla \theta_m, a_k) = (f_1, a_k), \\
(\partial_t S_m, a_k) + \frac{1}{Le}(\nabla S_m, \nabla a_k) + (\mathbf{u}_m \cdot \nabla S_m, a_k) = \frac{\alpha_*}{Le}(\nabla \theta_m, \nabla a_k) + (f_2, a_k),
\end{cases}
\tag{2.8}
$$

$$
\mathbf{u}_m(0) = \mathbf{u}_{0m}, \quad \theta_m(0) = \theta_{0m} \quad \text{and} \quad S_m(0) = S_{0m}, \quad k = 1, \ldots, m,
$$

where $(\mathbf{u}_{0m}, \theta_{0m}, S_{0m})$ is the L^2 orthogonal projection of $(\mathbf{u}_0, \theta_0, S_0)$ onto the space $\mathbf{V}^m$. Since (2.8) is an initial value problem for nonlinear ODEs, existence of unique local solutions in some neighborhood $[0, t_m)$, for some $t_m > 0$, follows by Picard–Lindelöf theorem. The *a priori* estimates we prove later in $L^\infty(L^2(\Omega))$-norm show that continuation of solutions beyond t_m follows. We will employ energy methods to derive *a priori* estimates. We first multiply $(2.8)_1$ by $c_k^{(m)}$, $(2.8)_2$ by $d_k^{(m)}$ and $(2.8)_3$ by $r_k^{(m)}$, and add these equations for $k = 1, \ldots, m$.

Using the skew-symmetry of the trilinear forms, we get

$$\begin{cases} \dfrac{1}{2}\dfrac{d}{dt}\|\mathbf{u}_m\|^2 + Pr\|\nabla\mathbf{u}_m\|^2 + \dfrac{1}{\epsilon}\|\mathbf{u}_m\|_{0,\Gamma}^2 = (Pr^2 Gr_\theta(\theta_m + N_r S_m)\mathbf{i}_3, \mathbf{u}_m) + \dfrac{1}{\epsilon}(\mathbf{g}, \mathbf{u}_m)_\Gamma, \\[2mm] \qquad\qquad \dfrac{1}{2}\dfrac{d}{dt}\|\theta_m\|^2 + \|\nabla\theta_m\|^2 = (f_1, \theta_m), \\[2mm] \qquad\qquad \dfrac{1}{2}\dfrac{d}{dt}\|S_m\|^2 + \dfrac{1}{Le}\|\nabla S_m\|^2 = \dfrac{\alpha_*}{Le}(\nabla\theta_m, \nabla S_m) + (f_2, S_m). \end{cases} \tag{2.9}$$

From the equation $(2.9)_2$, we readily obtain the *a priori* estimate

$$\sup_{t\in[0,T]} \|\theta_m(t)\|^2 + \|\theta_m\|_{L^2(H_0^1(\Omega))}^2 \le \|\theta_0\|^2 + \|f_1\|_{L^2(H^{-1}(\Omega))}^2 =: M_1. \tag{2.10}$$

From $(2.9)_3$, we obtain

$$\frac{d}{dt}\|S_m\|^2 + \frac{1}{Le}\|\nabla S_m\|^2 \le \frac{2\alpha_*}{Le}\|\nabla\theta_m\|^2 + 2Le\|f_2\|_{-1}^2.$$

Integrating this with respect time and using the fact that θ_m remains bounded in a bounded set of $L^2(0,T;H_0^1(\Omega))$, we obtain

$$\sup_{t\in[0,T]} \|S_m(t)\|^2 + \frac{1}{Le}\|S_m\|_{L^2(H_0^1(\Omega))}^2 \le \|S_0\|^2 + 2Le\|f_2\|_{L^2(H^{-1})} + \frac{2\alpha_*}{Le}M_1 =: M_2. \tag{2.11}$$

Let us now turn to the *a priori* estimate for $\mathbf{u}_m$. First we note that the right hand side of $(2.9)_1$ can be majorized using Young's inequality as follows

$$(Pr^2 Gr_\theta(\theta_m + N_r S_m)\mathbf{i}_3, \mathbf{u}_m) + \frac{1}{\epsilon}(\mathbf{g}, \mathbf{u}_m)_\Gamma \le \frac{Pr^3 Gr^2}{2\lambda}(\|\theta_m\|^2 + N_r^2\|S_m\|^2) + \frac{\lambda Pr}{2}\|\mathbf{u}_m\|^2$$
$$+ \frac{1}{\epsilon}\|\mathbf{g}\|_{0,\Gamma}^2 + \frac{1}{4\epsilon}\|\mathbf{u}_m\|_{0,\Gamma}^2.$$

Moreover, by the assumption that $\epsilon \in (0, \frac{1}{Pr})$ and Poincaré–Friedrich's inequality, we have

$$Pr\|\nabla\mathbf{u}_m\|^2 + \frac{1}{\epsilon}\|\mathbf{u}_m\|_{0,\Gamma}^2 \ge \frac{Pr}{2}\|\nabla\mathbf{u}_m\|^2 + \frac{1}{2\epsilon}\|\mathbf{u}_m\|_{0,\Gamma}^2 + \frac{\lambda Pr}{2}\|\mathbf{u}_m\|^2.$$

Employing these in $(2.9)_1$ we obtain

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u}_m\|^2 + \frac{Pr}{2}\|\nabla\mathbf{u}_m\|^2 + \frac{1}{4\epsilon}\|\mathbf{u}_m\|_{0,\Gamma}^2 \le \frac{Pr^3 Gr^2}{2\lambda}(\|\theta_m\|^2 + N_r^2\|S_m\|^2) + \frac{1}{\epsilon}\|\mathbf{g}\|_{0,\Gamma}^2.$$

Integrating this with respect to time and using (2.10)–(2.11), we obtain

$$\sup_{t\in[0,T]} \|\mathbf{u}_m\|^2 + Pr\|\nabla\mathbf{u}_m\|_{L^2(\mathbf{H})}^2 + \frac{1}{2\epsilon}\|\mathbf{u}_m\|_{L^2(L^2(\Gamma))}^2 \le M_3 + \frac{2}{\epsilon}\|\mathbf{g}\|_{L^2(L^2(\Gamma))}^2. \tag{2.12}$$

This last *a priori* bound can be used to obtain bounds for $\partial_t\mathbf{u}_m$. In fact, for $\mathbf{v}\in\mathbb{V}$, $c(\mathbf{u}_m, \mathbf{u}_m, \mathbf{v}) = (\mathbf{u}_m\cdot\nabla\mathbf{v}, \mathbf{u}_m)$ and thus by (A.6),

$$\sup_{\mathbf{v}\in\mathbb{V}} \frac{|(\partial_t\mathbf{u}_m, \mathbf{v})|}{\|\mathbf{v}\|_1} \le Pr\|\nabla\mathbf{u}_m\| + \sqrt{2}\|\mathbf{u}_m\|\|\nabla\mathbf{u}_m\| + Pr^2 Gr_\theta(\|\theta_m\| + N_r\|S_m\|).$$

Therefore $\|\partial_t\mathbf{u}_m\|_{L^2(\mathbb{V}^*)}$ is bounded. Similarly, for $\mathbf{v}\in\mathbf{V}$, by (A-6),

$$\sup_{\mathbf{v}\in\mathbf{V}} \frac{|(\partial_t\mathbf{u}_m, \mathbf{v})|}{\|\mathbf{v}\|_1} \le Pr\|\nabla\mathbf{u}_m\| + C\|\mathbf{u}_m\|_1^2 + Pr^2 Gr_\theta(\|\theta_m\| + N_r\|S_m\|) + \frac{1}{\epsilon}\|\mathbf{g}_m\|_{L^2(\Gamma)} + \frac{1}{\epsilon}\|\mathbf{u}_m\|_{L^2(\Gamma)}$$

and thus $\|\partial_t \mathbf{u}_m\|_{L^1(\mathbf{V}^*)}$ is bounded due to the bounds in (2.10)−(2.12). Similarly, we can show $\|\partial_t \theta_m\|_{L^2(H^{-1}(\Omega))}$ and $\|\partial_t S_m\|_{L^2(H^{-1}(\Omega))}$ are bounded as well.

The *a priori* estimates we obtained so far allows us to extract subsequences again denoted by $\{(\mathbf{u}_m, \theta_m, S_m)\}_{n=1}^\infty$ such that

$$\mathbf{u}_m \to \mathbf{u} \begin{cases} \text{weakly in } L^2(0,T;\mathbf{V}) \\ \text{weak star in } L^\infty(0,T;\mathbf{H}) \\ \text{strongly in } L^2(0,T;\mathbf{H}), \end{cases} \quad \text{and} \quad (\theta_m, S_m) \to (\theta, S) \begin{cases} \text{weakly in } L^2(0,T;H_0^1(\Omega)) \\ \text{weak star in } L^\infty(0,T;L^2(\Omega)) \\ \text{strongly in } L^2(0,T;L^2(\Omega)). \end{cases}$$

Here the strong convergence follows by the Aubin–Simon compactness Lemma ([27], Cor. 4, p. 85), as we have the imbeddings $\mathbf{V} \hookrightarrow\hookrightarrow \mathbf{H} \hookrightarrow \mathbf{V}^*$ and $H_0^1(\Omega) \hookrightarrow\hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$.

Moreover,

$$\partial_t \mathbf{u}_m \to \partial_t \mathbf{u} \text{ weakly in } L^2(0,T;\mathbb{V}^*), \quad (\partial_t \theta_m, \partial_t S_m) \to (\partial_t \theta, \partial_t S) \text{ weakly in } L^2(0,T;H^{-1}(\Omega))$$

and $\mathbf{u}_m|_\Gamma \to \mathbf{u}|_\Gamma$ weakly in $L^2(L^2(\Gamma))$. When taking limit in (2.8), it is convenient for us to use the trilinear form $c(\cdot,\cdot,\cdot)$ involving boundary term, *i.e.*,

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

The presence of nonlinear boundary term requires us to also show that $\mathbf{u}_m \cdot \mathbf{n} \to \mathbf{u} \cdot \mathbf{n}$ in $L^2(0,T;L^2(\Gamma))$ strongly. In order to prove such a convergence, we first recall the integration by parts formula ([12], Eq. (I.2.17), p. 28).

$$\langle \mathbf{u}_m \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n}, \mathbf{v} \rangle_\Gamma = \int_\Omega \mathbf{v} \nabla \cdot (\mathbf{u}_m - \mathbf{u}) \, d\Omega + \int_\Omega (\mathbf{u}_m - \mathbf{u}) \cdot \nabla \mathbf{v} \, d\Omega \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega). \tag{2.13}$$

By solving the variational problem

$$\begin{cases} (\partial_t \mathbf{v}, \phi) + (\nabla \mathbf{v}, \nabla \phi) = 0 & \forall \phi \in H_0^1(\Omega) \\ \mathbf{v}|_\Gamma = \mathbf{u} \cdot \mathbf{n} - \mathbf{u}_m \cdot \mathbf{n} \\ \mathbf{v}(x,0) = 0 & \text{in } \Omega \end{cases} \tag{2.14}$$

we obtain a unique solution $\mathbf{v} \in L^2(0,T;\mathbf{H}^1(\Omega))$ such that

$$\|\nabla \mathbf{v}\|_{L^2(\mathbf{L}^2(\Omega))} \leq C\|\mathbf{u} \cdot \mathbf{n} - \mathbf{u}_m \cdot \mathbf{n}\|_{W_1^\Sigma},$$

see Lemma 3.4. Therefore by taking $\mathbf{v}$ in (2.13) to be the unique solution of this variational problem (2.14) and using the fact that $\nabla \cdot \mathbf{u}_m = \nabla \cdot \mathbf{u} = 0$ yields

$$\|\mathbf{u}_m \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n}\|_{L^2(L^2(\Gamma))} \leq \|\mathbf{u}_m - \mathbf{u}\|_{L^2(\mathbf{L}^2(\Omega))} \|\mathbf{u}_m \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n}\|_{W_1^\Sigma}. \tag{2.15}$$

The weak convergence $\mathbf{u}_m \to \mathbf{u}$ in W_1 and trace theorem imply that $\mathbf{u}_m \cdot \mathbf{n} \to \mathbf{u} \cdot \mathbf{n}$ weakly in W_1^Σ and thus $\|\mathbf{u}_m \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n}\|_{W_1^\Sigma}$ is bounded. Therefore the required strong convergence follows from (2.15).

Let $\psi_i(t)$, $i = 1, 2, 3$, be a continuously differentiable function on $[0, T]$ with $\psi_i(T) = 0$. We multiply $(2.8)_i$ by ψ_i, $i = 1, 2, 3$ and integrate with respect to time. Further, we integrate by parts in the time derivative term to move the derivative onto ψ_i. Now we can take limit in (2.8) by using standard techniques and show $(\mathbf{u}, \theta, S)$ is indeed a solution of (2.7). The *a priori* estimates in the lemma follow by taking the limit on the *a priori* estimates (2.10)−(2.12) and using the weak lower semi-continuity of the norms.

Before we prove the uniqueness of the solutions, let us introduce the continuous linear forms $B_1(\mathbf{u}, \mathbf{v}), B_2(\mathbf{u}, \mathbf{v}), B(\mathbf{u}, \mathbf{v}) \in \mathbf{V}^*$ defined by

$$\langle B_1(\mathbf{u}, \mathbf{v}), \mathbf{w} \rangle_{\mathbf{V}^*, \mathbf{V}} = (\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{w}), \quad \langle B_2(\mathbf{u}, \mathbf{v}), \mathbf{w} \rangle_{\mathbf{V}^*, \mathbf{V}} = (\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v})$$

and

$$\langle B(\mathbf{u}, \mathbf{v}), \mathbf{w}\rangle_{\mathbf{V}^*, \mathbf{V}} = c(\mathbf{u}, \mathbf{w}, \mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}.$$

Notice that by (A-6), $B_1(\mathbf{u}, \mathbf{u}) \in L^{4/3}(0, T; L^{4/3}(\Omega))$ and $B_2(\mathbf{u}, \mathbf{u}) \in L^2(0, T; \mathbf{V}^*)$ and thus $\mathbf{u}_t \in L^2(0, T; \mathbf{V}^*) + L^{4/3}(0, T; L^{4/3}(\Omega))$. Moreover, by *a priori* bounds and (2.1), $\mathbf{u} \in L^2(0, T; \mathbf{V}) \cap L^4(0, T; L^4(\Omega))$. Therefore we have $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2 = 2(\mathbf{u}'(t), \mathbf{u}(t))$ and similar identity holds for θ and S as well.

Assume $(\mathbf{u}, \theta, S)$ and $(\mathbf{u}_1, \theta_1, S_1)$ be two solutions of (2.7) and let $(\mathbf{w}, \hat{\theta}, \hat{S})$ be the difference between them. Then the difference satisfies

$$\begin{cases} (\partial_t \mathbf{w}, \mathbf{v}) + Pr(\nabla \mathbf{w}, \nabla \mathbf{v}) + c(\mathbf{w}, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{w}, \mathbf{v}) + \dfrac{1}{\epsilon}(\mathbf{w}, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\hat{\theta} + N_r \hat{S})\mathbf{i}_3, \mathbf{v}), \\[2mm] (\partial_t \hat{\theta}, \phi) + (\nabla \hat{\theta}, \nabla \phi) + (\mathbf{w} \cdot \nabla \theta, \phi) + (\mathbf{u} \cdot \nabla \hat{\theta}, \phi) = 0, \\[2mm] (\partial_t \hat{S}, \psi) + \dfrac{1}{Le}(\nabla \hat{S}, \nabla \psi) + (\mathbf{w} \cdot \nabla S, \psi) + (\mathbf{u} \cdot \nabla \hat{S}, \psi) = \dfrac{\alpha_*}{Le}(\nabla \hat{\theta}, \nabla \psi), \end{cases} \tag{2.16}$$

and

$$\mathbf{w}(\mathbf{x}, 0) = 0, \quad \hat{\theta}(\mathbf{x}, 0) = 0 \quad \text{and} \quad \hat{S}(\mathbf{x}, 0) = 0,$$

$\forall (\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$, for a.e. $t \in (0, T]$.

Setting $\phi = \hat{\theta}$ and $\psi = \hat{S}$ in $(2.16)_2$ and $(2.16)_3$, respectively, using the skew-symmetry of the tri-linear forms and estimating the tri-linear terms $(\mathbf{w} \cdot \nabla \theta, \hat{\theta})$ and $(\mathbf{w} \cdot \nabla S, \hat{S})$ using (A-1) yields

$$\begin{cases} \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\|\hat{\theta}\|^2 + \dfrac{1}{2}\|\nabla \hat{\theta}\|^2 \leq \dfrac{Pr}{8\Lambda}\|\nabla \mathbf{w}\|^2 + C\|\mathbf{w}\|^2\|\theta\|_{L^4(\Omega)}^4 \\[2mm] \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\|\hat{S}\|^2 + \dfrac{1}{2Le}\|\nabla \hat{S}\|^2 \leq \dfrac{Pr}{8}\|\nabla \mathbf{w}\|^2 + C\|\mathbf{w}\|^2\|S\|_{L^4(\Omega)}^4 + \dfrac{\alpha_*}{Le}\|\nabla \hat{\theta}\|\|\nabla \hat{S}\|. \end{cases} \tag{2.17}$$

Multiplying $(2.17)_1$ by Λ and adding the result to $(2.17)_2$ yields

$$\begin{aligned} \dfrac{\mathrm{d}}{\mathrm{d}t}(\Lambda\|\hat{\theta}\|^2 + \|\hat{S}\|^2) + \Lambda\|\nabla \hat{\theta}\|^2 + \|\nabla \hat{S}\|^2 &\leq \dfrac{Pr}{4}\|\nabla \mathbf{w}\|^2 + C\|\mathbf{w}\|^2 \left(\|\theta\|_{L^4(\Omega)}^4 + \|S\|_{L^4(\Omega)}^4\right) \\ &\quad + \dfrac{2\alpha_*}{Le}\|\nabla \hat{\theta}\|\|\nabla \hat{S}\|. \end{aligned} \tag{2.18}$$

Let us assume $\alpha_* > 0$ without loss of generality and use Young's inequality to write

$$\dfrac{-2\alpha_*}{Le}\|\nabla \hat{\theta}\|\|\nabla \hat{S}\| \geq -\dfrac{\alpha_*}{2\epsilon Le}\|\nabla \hat{\theta}\|^2 - \dfrac{\epsilon \alpha_*}{2Le}\|\nabla \hat{S}\|^2$$

for some $\epsilon > 0$. Therefore

$$\Lambda\|\nabla \hat{\theta}\|^2 + \dfrac{1}{Le}\|\nabla \hat{S}\|^2 - \dfrac{2\alpha_*}{Le}\|\nabla \hat{\theta}\|\|\nabla \hat{S}\| \geq \left(\Lambda - \dfrac{\alpha_*}{2\epsilon Le}\right)\|\nabla \hat{\theta}\|^2 + \left(\dfrac{1}{Le} - \dfrac{\alpha_* \epsilon}{2Le}\right)\|\nabla \hat{S}\|^2.$$

Now we can choose ϵ such that

$$\Lambda\|\nabla \hat{\theta}\|^2 + \dfrac{1}{Le}\|\nabla \hat{S}\|^2 - \dfrac{2\alpha_*}{Le}\|\nabla \hat{\theta}\|\|\nabla \hat{S}\| \geq \lambda_1\|\nabla \hat{\theta}\|^2 + \lambda_2\|\nabla \hat{S}\|^2$$

for some positive numbers λ_1 and λ_2. Thus from (2.18), we obtain

$$\dfrac{\mathrm{d}}{\mathrm{d}t}(\Lambda\|\hat{\theta}\|^2 + \|\hat{S}\|^2) + \lambda_1\|\nabla \hat{\theta}\|^2 + \lambda_2\|\nabla \hat{S}\|^2 \leq \dfrac{Pr}{4}\|\nabla \mathbf{w}\|^2 + C\|\mathbf{w}\|^2 \left(\|\theta\|_{L^4(\Omega)}^4 + \|S\|_{L^4(\Omega)}^4\right). \tag{2.19}$$

Similarly setting $\mathbf{v} = \mathbf{w}$ in $(2.16)_1$ and estimating the term $c(\mathbf{w}, \mathbf{u}, \mathbf{w})$ using (A-2), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{w}\|^2 + \frac{Pr}{2}\|\nabla\mathbf{w}\|^2 + \frac{1}{2\epsilon}\|\mathbf{w}\|_{0,\Gamma}^2 \leq C(\|\nabla\mathbf{u}\|^2 + \|\mathbf{u}\|_{L^4(\Omega)}^4)\|\mathbf{w}\|^2 + C\left(\|\hat{\theta}\|^2 + \|\hat{S}\|^2\right). \tag{2.20}$$

Therefore adding (2.19) and (2.20), we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|(\mathbf{w}, \Lambda\hat{\theta}, \hat{S})\|^2 + \frac{Pr}{2}\|\nabla\mathbf{w}\|^2 + \lambda_1\|\nabla\hat{\theta}\|^2 + \lambda_2\|\nabla\hat{S}\|^2 + \frac{1}{\epsilon}\|\mathbf{w}\|_{0,\Gamma}^2 \leq \hat{G}(t)\|(\mathbf{w}, \Lambda\hat{\theta}, \hat{S})\|^2 \tag{2.21}$$

where $\hat{G}(t) := C[\|\theta\|_{L^4(\Omega)}^4 + \|S\|_{L^4(\Omega)}^4 + \|\nabla\mathbf{u}\|^2 + \|\mathbf{u}\|_{L^4(\Omega)}^4]$. Notice that, by the *a priori* estimates and (2.1), $\hat{G}(t) \in L^1(0,T)$ and an application of Gronwall's lemma to (2.21) yields $\|(\mathbf{w}, \hat{\theta}, \hat{S})\| = 0$. This proves uniqueness. $\qquad\square$

3. PENALIZED BOUNDARY CONTROL PROBLEM

In this section, we formulate the penalized boundary control problem and prove the convergence of its solutions to the solutions of Dirichlet control problem as the penalty parameter ϵ goes to zero.

The optimal boundary control problem we will study is the following: find $(\mathbf{u}, \theta, S, \mathbf{g})$ such that the cost functional

$$\mathcal{J}(\mathbf{u}, \mathbf{g}) = \int_0^T \Theta(\mathbf{u})\,\mathrm{d}t + \frac{\gamma}{2}\int_0^T \int_\Gamma |\mathbf{g}|^2 \,\mathrm{d}\Gamma\,\mathrm{d}t \tag{3.1}$$

is minimized subject to (2.2), where $\mathbf{g}$ is the control field and γ is a positive parameter. In this paper, we assume the function $\Theta : \mathbf{H}^1(\Omega) \to \mathbb{R}^+$ appearing in the cost functional is assumed to satisfy the following:

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

The allowed class of functionals of course includes regulation of velocity gradient, *i.e.*, $\Theta(\mathbf{u}) = \delta\|\nabla\mathbf{u}\|^2$. The other allowed functionals include regulation of kinetic energy in weighted H^1- norm, *i.e.*, $\Theta(\mathbf{u}) = \frac{\delta_1}{2}\|\nabla\mathbf{u}\|^2 + \frac{\delta_2}{2}\|\mathbf{u}\|^2$ and regulation of square of the vorticity, *i.e.*, $\Theta(\mathbf{u}) = \frac{\delta}{2}\|\nabla\times\mathbf{u}\|^2$. The former expression for Θ already satisfies Assumption (B). The later can also be shown to satisfy Assumption (B), see Lemma 3.3 of [17].

Letting $\mathcal{U}_{\mathrm{ad}} := \mathbf{W}_1^\Sigma$ and $\mathcal{U}_{\mathrm{ad}}^\epsilon := L^2(L^2(\Gamma))$ be the admissible control sets for the Dirichlet and penalized control problems, respectively, we formulate the optimal control problems as follows.

Dirichlet boundary control.
(P)   find a control $\mathbf{g} \in \mathcal{U}_{\mathrm{ad}}$ *such that the cost functional* (3.1) *is minimized subject to* (2.5).

Penalized boundary control.
(P_ϵ)   For each $\epsilon > 0$, *find a control* $\mathbf{g}_\epsilon \in \mathcal{U}_{\mathrm{ad}}^\epsilon$ *such that the cost functional* (3.1) *is minimized subject to* (2.7).

Let $\mathcal{A}_{\mathrm{ad}}^\epsilon$ and $\mathcal{A}_{\mathrm{ad}}$ be the set of all admissible states and controls for the control problems $(P)_\epsilon$ and (P), respectively. That is,

$$\mathcal{A}_{\mathrm{ad}}^\epsilon := \{(\mathbf{u}, \theta, S, \mathbf{g}) \in \mathcal{W}(0,T) \times \mathcal{U}_{\mathrm{ad}}^\epsilon \,|\, \mathcal{J}(\mathbf{u}, \mathbf{g}) < \infty \text{ and } (2.7) \text{ is satisfied}\}$$

and

$$\mathcal{A}_{\mathrm{ad}} := \{(\mathbf{u}, \theta, S, \mathbf{g}) \in \mathcal{W}(0,T) \times \mathcal{U}_{\mathrm{ad}} \,|\, \mathcal{J}(\mathbf{u}, \mathbf{g}) < \infty \text{ and } (2.5) \text{ is satisfied}\}.$$

Theorem 3.1. *Assume* $\epsilon \in (0, \frac{1}{Pr})$. *Then there exists an optimal solution* $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon, \mathbf{g}_\epsilon) \in \mathcal{A}_{\mathrm{ad}}^\epsilon$ *that minimizes* (3.1) *subject to* (2.7).

Proof. Since it follows from Lemma 2.3 that there exists a solution $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon, \mathbf{g}_\epsilon) \in \mathcal{W}(0,T) \times L^2(0,T;L^2(\Gamma))$ such that (2.7) holds, $\mathcal{A}_{\mathrm{ad}}^\epsilon$ is non empty. Let $\{(\mathbf{u}_n, \theta_n, S_n, \mathbf{g}_n)\}$ be a minimizing sequence in $\mathcal{A}_{\mathrm{ad}}^\epsilon$, *i.e.*,

$$\lim_{n \to \infty} \mathcal{J}(\mathbf{u}_n, \mathbf{g}_n) = \inf \left\{ \mathcal{J}(\mathbf{u}, \mathbf{g}) : (\mathbf{u}, \theta, S, \mathbf{g}) \in \mathcal{A}_{\mathrm{ad}}^\epsilon \right\},$$

where $(\mathbf{u}_n, \theta_n, S_n)$ satisfy

$$\begin{cases} (\partial_t \mathbf{u}_n, \mathbf{v}) + Pr(\nabla \mathbf{u}_n, \nabla \mathbf{v}) + c(\mathbf{u}_n, \mathbf{u}_n, \mathbf{v}) + \dfrac{1}{\epsilon}(\mathbf{u}_n, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\theta_n + N_r S_n)\mathbf{i}_3, \mathbf{v}) \\ \qquad\qquad\qquad + \dfrac{1}{\epsilon}(\mathbf{g}_n, \mathbf{v})_\Gamma, \quad \forall \mathbf{v} \in \mathbf{V}(\Omega) \\ (\partial_t \theta_n, \phi) + (\nabla \theta_n, \nabla \phi) + (\mathbf{u}_n \cdot \nabla \theta_n, \phi) = 0, \quad \forall \phi \in H_0^1(\Omega) \\ (\partial_t S_n, \psi) + \dfrac{1}{Le}(\nabla S_n, \nabla \psi) + (\mathbf{u}_n \cdot \nabla S_n, \psi) = \dfrac{\alpha^*}{Le}(\nabla \theta, \nabla \psi), \quad \forall \psi \in H_0^1(\Omega). \end{cases} \quad (3.2)$$

The boundedness of $\{\mathcal{J}(\mathbf{u}_n, \mathbf{g}_n)\}$ in $\mathbb{R}^+$ implies that $\|\mathbf{g}_n\|_{L^2(0,T;L^2(\Gamma))} \le C_g$ for all n. The *a priori* estimates in Lemma 2.3 now allows us to extract subsequences still denoted by $\{(\mathbf{u}_n, \theta_n, S_n, \mathbf{g}_n)\}$, such that

$$\{\mathbf{g}_n, \mathbf{u}_n, \theta_n, S_n, \partial_t \theta_n, \partial_t S_n, \mathbf{u}_n|_\Gamma\} \rightharpoonup \{\mathbf{g}^*, \mathbf{u}^*, \theta^*, S^*, \partial_t \theta^*, \partial_t S^*, \mathbf{u}^*|_\Gamma\}$$

weakly in $\quad L^2(L^2(\Gamma)) \times L^2(\mathbf{V}) \times (L^2(H_0^1(\Omega)))^2 \times (L^2(H^{-1}(\Omega))^2 \times L^2(L^2(\Gamma)).$

Moreover, as $(\partial_t \mathbf{u}_n, \partial_t \theta_n, \partial_t S_n)$ is bounded in $L^1(0,T; \mathbf{V}^*) \times L^2(0,T;L^2(\Omega))^2$, by Aubin–Simon's compactness Lemma ([27], Cor. 4, p. 85), we have

$$(\mathbf{u}_n, \theta_n, S_n) \to (\mathbf{u}^*, \theta^*, S^*) \quad \text{strongly in} \quad L^2(0,T;\mathbf{H}) \times L^2(0,T;L^2(\Omega))^2.$$

These convergence relations allow us to pass to the limit in (3.2) to show that $(\mathbf{u}^*, \theta^*, S^*, \mathbf{g}^*)$ satisfies (2.7). Moreover, from the uniqueness of the weak solutions, we conclude that $(\mathbf{u}^*, \theta^*, S^*, \mathbf{g}^*) \in \mathcal{A}_{\mathrm{ad}}^\epsilon$. Finally, using the sequential weak lower semi-continuity of the functional $\mathcal{J}$, we have

$$\mathcal{J}(\mathbf{u}^*, \mathbf{g}^*) \le \liminf_{n \to \infty} \mathcal{J}(\mathbf{u}_n, \mathbf{g}_n)$$
$$= \inf\{\mathcal{J}(\mathbf{u}, \mathbf{g}) : (\mathbf{u}, \theta, S, \mathbf{g}) \in \mathcal{A}_{\mathrm{ad}}^\epsilon\}.$$

This shows $(\mathbf{u}^*, \theta^*, S^*, \mathbf{g}^*)$ is indeed an optimal solution. $\qquad\square$

3.1. Convergence of solutions of penalized control problem

In this section, we will study the convergence of solutions $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon)$ of penalized optimal control problem $(P)_\epsilon$ as $\epsilon \to 0$.

Theorem 3.2. *For each $\epsilon \in (0, \frac{1}{Pr})$, let $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon, \mathbf{g}_\epsilon) \in W(0,T) \times \mathcal{U}_{\mathrm{ad}}^\epsilon$ be a solution to optimal penalized control problem $(P)_\epsilon$. Then there exists a $(\widehat{\mathbf{u}}, \widehat{\theta}, \widehat{S}, \widehat{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$ and a subsequence $\{\epsilon_k\}_{k=1}^\infty$ such that as $k \to \infty$, $\mathbf{u}_{\epsilon_k} \rightharpoonup \widehat{\mathbf{u}}$ in $L^2(\mathbf{V})$, $(\theta_{\epsilon_k}, S_{\epsilon_k}) \rightharpoonup (\widehat{\theta}, \widehat{S})$ in $L^2(H^1(\Omega))^2$ and $\mathbf{g}_{\epsilon_k} \rightharpoonup \widehat{\mathbf{g}}$ in $\mathcal{U}_{\mathrm{ad}}^\epsilon$. Moreover, $(\mathbf{u}_{\epsilon_k}, \theta_{\epsilon_k}, S_{\epsilon_k}) \to (\widehat{\mathbf{u}}, \widehat{\theta}, \widehat{S})$ in $L^2(\mathbf{H}) \times L^2(L^2(\Omega))^2$.*

Proof. Let us first obtain some bounds for $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon, \mathbf{g}_\epsilon)$ independent of ϵ. Let $(\widetilde{\mathbf{u}}_\epsilon, \widetilde{\theta}_\epsilon, \widetilde{S}_\epsilon)$ be the solution of (2.7) with $\mathbf{g} = 0$. Then it satisfies

$$\begin{cases} (\partial_t \widetilde{\mathbf{u}}_\epsilon, \mathbf{v}) + Pr(\nabla \widetilde{\mathbf{u}}_\epsilon, \nabla \mathbf{v}) + c(\widetilde{\mathbf{u}}_\epsilon, \widetilde{\mathbf{u}}_\epsilon, \mathbf{v}) + \dfrac{1}{\epsilon}(\widetilde{\mathbf{u}}_\epsilon, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\widetilde{\theta}_\epsilon + N_r \widetilde{S}_\epsilon)\mathbf{i}_3, \mathbf{v}), \\ \qquad\qquad (\partial_t \widetilde{\theta}_\epsilon, \phi) + (\nabla \widetilde{\theta}_\epsilon, \nabla \phi) + (\widetilde{\mathbf{u}}_\epsilon \cdot \nabla \widetilde{\theta}_\epsilon, \phi) = (f_1, \phi), \\ \qquad\qquad (\partial_t \widetilde{S}_\epsilon, \psi) + \dfrac{1}{Le}(\nabla \widetilde{S}_\epsilon, \nabla \psi) + (\widetilde{\mathbf{u}}_\epsilon \cdot \nabla \widetilde{S}_\epsilon, \psi) = \dfrac{\alpha_*}{Le}(\nabla \widetilde{\theta}_\epsilon, \nabla \psi) + (f_2, \psi), \end{cases} \quad (3.3)$$

and

$$\widetilde{\mathbf{u}}_\epsilon(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \widetilde{\theta}_\epsilon(\mathbf{x}, 0) = \theta_0(\mathbf{x}) \quad \text{and} \quad \widetilde{S}_\epsilon(\mathbf{x}, 0) = S_0(\mathbf{x}),$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$. By setting $\mathbf{v} = \widetilde{\mathbf{u}}_\epsilon$ and arguing as the proof of Lemma 2.3, we obtain

$$\sup_{t \in [0,T]} \|\widetilde{\mathbf{u}}_\epsilon\|^2 + Pr\|\nabla \widetilde{\mathbf{u}}_\epsilon\|_{L^2(\mathbf{H})}^2 + \frac{1}{2\epsilon}\|\widetilde{\mathbf{u}}_\epsilon\|_{L^2(L^2(\Gamma))}^2 \leq M_3, \tag{3.4}$$

where M_3 is as defined in Lemma 2.3. Since $(\widetilde{\mathbf{u}}_\epsilon, \widetilde{\theta}_\epsilon, \widetilde{S}_\epsilon, \mathbf{0}) \in \mathcal{A}_{\text{ad}}^\epsilon$, we have

$$\mathcal{J}(\mathbf{u}_\epsilon, \mathbf{g}_\epsilon) \leq \mathcal{J}(\widetilde{\mathbf{u}}_\epsilon, 0).$$

Therefore by Assumption (B) and (3.4), we have

$$\int_0^T \Theta(\mathbf{u}_\epsilon) \, \mathrm{d}t + \frac{\gamma}{2}\|\mathbf{g}_\epsilon\|_{L^2(L^2(\Gamma))}^2 \leq \hat{c}_1\|\nabla \widetilde{\mathbf{u}}_\epsilon\|_{L^2(\mathbf{H})}^2 + \hat{c}_2\|\widetilde{\mathbf{u}}_\epsilon\|_{L^2(L^2(\Gamma))}^2 = \frac{M_3}{Pr}(\hat{c}_1 + 2\hat{c}_2)$$

which yields

$$\int_0^T \Theta(\mathbf{u}_\epsilon) \, \mathrm{d}t < M_4 \quad \text{and} \quad \|\mathbf{g}_\epsilon\|_{L^2(L^2(\Gamma))}^2 < \frac{2M_4}{\gamma},$$

where $M_4 := \frac{M_3}{Pr}(\hat{c}_1 + 2\hat{c}_2)$. Moreover, by Assumption (B) and *a priori* estimates in Lemma 2.3 yield

$$c_1\|\nabla \mathbf{u}_\epsilon\|_{L^2(\mathbf{H})}^2 \leq M_4 + c_2\|\mathbf{u}_\epsilon\|_{L^2(L^2(\Gamma))}^2 \leq M_4 + c_2 \left[\frac{2M_3}{Pr} + \frac{8M_4}{\gamma}\right],$$

which shows $\|\nabla \mathbf{u}_\epsilon\|_{L^2(\mathbf{H})}$ is bounded independent of ϵ. By Poincaré−Friedrichs' inequality, we also have a bound on $\|\mathbf{u}_\epsilon\|_{L^2(\mathbf{H})}$ independent of ϵ. On the other hand by the *a priori* estimates in Lemma 2.3, we have that $\|\theta_\epsilon\|_{L^\infty(L^2(\Omega)) \cap L^2(H_0^1(\Omega))}$ and $\|S_\epsilon\|_{L^\infty(L^2(\Omega)) \cap L^2(H_0^1(\Omega))}$ are bounded independent of ϵ. Now notice that for each $(\mathbf{v}, \phi, \psi) \in \mathbb{V} \times (H_0^1(\Omega))^2$ equation (2.7) reduces to

$$\begin{cases} (\partial_t \mathbf{u}_\epsilon, \mathbf{v}) + Pr(\nabla \mathbf{u}_\epsilon, \nabla \mathbf{v}) + ((\mathbf{u}_\epsilon \cdot \nabla)\mathbf{u}_\epsilon, \mathbf{v}) = (Pr^2 Gr_\theta(\theta_\epsilon + N_r S_\epsilon)\mathbf{i}_3, \mathbf{v}), \\ (\partial_t \theta_\epsilon, \phi) + (\nabla \theta_\epsilon, \nabla \phi) + ((\mathbf{u}_\epsilon \cdot \nabla)\theta_\epsilon, \phi) = (f_1, \phi), \\ (\partial_t S_\epsilon, \psi) + \frac{1}{Le}(\nabla S_\epsilon, \nabla \psi) + ((\mathbf{u}_\epsilon \cdot \nabla)S_\epsilon, \psi) = \frac{\alpha_*}{Le}(\nabla \theta_\epsilon, \nabla \psi) + (f_2, \psi), \end{cases} \tag{3.5}$$

and

$$\mathbf{u}_\epsilon(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \theta_\epsilon(\mathbf{x}, 0) = \theta_0(\mathbf{x}) \quad \text{and} \quad S_\epsilon(\mathbf{x}, 0) = S_0(\mathbf{x}).$$

From $(3.5)_1$, we easily have that

$$\sup_{\mathbf{v} \in \mathbb{V}} \frac{|(\partial_t \mathbf{u}_\epsilon, \mathbf{v})|}{\|\mathbf{v}\|_1} \leq Pr\|\nabla \mathbf{u}_\epsilon\| + \sqrt{2}\|\mathbf{u}_\epsilon\|\|\nabla \mathbf{u}_\epsilon\| + Pr^2 Gr_\theta(\|\theta_\epsilon\| + Nr\|S_\epsilon\|).$$

Thus

$$\|\partial_t \mathbf{u}_\epsilon\|_{L^1(\mathbb{V}^*)} \leq C[\|\nabla \mathbf{u}_\epsilon\|_{L^2(\mathbf{H})} + \|\mathbf{u}_\epsilon\|_{L^2(\mathbf{H})} + \|\theta_\epsilon\|_{L^\infty(L^2(\Omega))} + \|S_\epsilon\|_{L^\infty(L^2(\Omega))}].$$

It now follows from the previously obtained bounds (ϵ independent) that $\|\partial_t \mathbf{u}_\epsilon\|_{L^1(\mathbb{V}^*)}$ is bounded independent of ϵ. Similarly, we can show that $\|\partial_t \theta_\epsilon\|_{L^1(H^{-1})}$ and $\|\partial_t S_\epsilon\|_{L^1(H^{-1})}$ are also bounded independent of ϵ.

These *a priori* bounds (ϵ independent) allow us to extract a subsequence $\{\mathbf{u}_{\epsilon_k}, \theta_{\epsilon_k}, S_{\epsilon_k}\}$ such that as $k \to \infty$,

$$\epsilon_k \to 0, \quad \mathbf{u}_{\epsilon_k} \rightharpoonup \hat{\mathbf{u}} \text{ in } L^2(\mathbf{V}), \quad \theta_{\epsilon_k} \rightharpoonup \hat{\theta} \text{ in } L^2(H^1(\Omega)), \quad S_{\epsilon_k} \rightharpoonup \hat{S} \text{ in } L^2(H^1(\Omega)).$$

Moreover, by Aubin–Simon compactness Lemma ([27], Cor. 4, p. 85), we have the following strong convergence

$$\mathbf{u}_{\epsilon_k} \to \hat{\mathbf{u}} \text{ in } L^2(\mathbf{H}), \quad \theta_{\epsilon_k} \to \hat{\theta} \text{ in } L^2(L^2(\Omega)), \quad S_{\epsilon_k} \to \hat{S} \text{ in } L^2(L^2(\Omega)).$$

Let $\psi_i(t)$, $i = 1, 2, 3$, be a continuously differentiable function on $[0, T]$ with $\psi_i(T) = 0$. We multiply $(3.5)_i$ by ψ_i, $i = 1, 2, 3$, integrate with respect to time and then integrate by parts in the time derivative term to move the derivative onto ψ_i. Now setting $\epsilon = \epsilon_k$ and letting $k \to \infty$ yields

$$\begin{cases} (\partial_t \hat{\mathbf{u}}, \mathbf{v}) + Pr(\nabla \hat{\mathbf{u}}, \nabla \mathbf{v}) + ((\hat{\mathbf{u}} \cdot \nabla)\hat{\mathbf{u}}, \mathbf{v}) = (Pr^2 Gr_\theta(\hat{\theta} + N_r \hat{S})\mathbf{i}_3, \mathbf{v}), \\ (\partial_t \hat{\theta}, \phi) + (\nabla \hat{\theta}, \nabla \phi) + ((\hat{\mathbf{u}} \cdot \nabla)\hat{\theta}, \phi) = (f_1, \phi), \\ (\partial_t \hat{S}, \psi) + \dfrac{1}{Le}(\nabla \hat{S}, \nabla \psi) + ((\hat{\mathbf{u}} \cdot \nabla)\hat{S}, \psi) = \dfrac{\alpha_*}{Le}(\nabla \hat{\theta}, \nabla \psi) + (f_2, \psi), \end{cases} \tag{3.6}$$

and

$$\hat{\mathbf{u}}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \hat{\theta}(\mathbf{x}, 0) = \theta_0(\mathbf{x}) \quad \text{and} \quad \hat{S}(\mathbf{x}, 0) = S_0(\mathbf{x}),$$

Let $\psi_1(t)$ be a infinitely continuously differentiable function with compact support in $[0, T]$. We multiply $(2.7)_1$ by ψ_1 and integrate with respect to time. Further, we integrate by parts in the time derivative term to move the derivative onto ψ_1. Now multiplying the resulting equation with $\epsilon = \epsilon_k$ by ϵ_k and letting $k \to \infty$ yields

$$\int_0^T \int_\Gamma \hat{\mathbf{u}} \cdot \mathbf{v}\psi_1 \, \mathrm{d}s\, \mathrm{d}t = \int_0^T \int_\Gamma \hat{\mathbf{g}} \cdot \mathbf{v}\psi_1 \, \mathrm{d}s\, \mathrm{d}t \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)$$

which implies $\hat{\mathbf{u}}|_\Gamma = \hat{\mathbf{g}}$. This proves that $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S}, \hat{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$. $\qquad\square$

We would like to show that the limit $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S}, \hat{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$ obtained in Theorem 3.2 is indeed an optimal solution to the Dirichlet control problem. That is we would like to show that

$$\mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}) \le \mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) \quad \forall (\check{\mathbf{u}}, \check{\theta}, \check{S}, \check{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}. \tag{3.7}$$

In the Lemma below, we prove the sub-optimality of the limit obtained in Theorem 3.2, in the sense that we have optimality within the class of functions in $\mathcal{A}_{\mathrm{ad}}$ that satisfy the additional regularity assumption

$$-\check{p}\mathbf{n} + Pr\frac{\partial \check{\mathbf{u}}}{\partial \mathbf{n}} - \frac{1}{2}(\check{\mathbf{u}} \cdot \mathbf{n})\check{\mathbf{u}} \in L^2(L^2(\Gamma)). \tag{3.8}$$

Lemma 3.3. *Assume $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S}, \hat{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$ is the limit defined in Theorem 3.2. Then*

$$\mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}) \le \mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) \quad \forall (\check{\mathbf{u}}, \check{\theta}, \check{S}, \check{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$$

satisfying (3.8).

Proof. As a proof of this theorem can be established using arguments used in the proof of a similar result in [17], we will only provide a sketch. Let $(\check{\mathbf{u}}, \check{\theta}, \check{S}, \check{\mathbf{g}})$ be an arbitrary element in $\mathcal{A}_{\mathrm{ad}}$ satisfying (3.8). Defining

$$\check{\mathbf{g}}_\epsilon = \check{\mathbf{g}} + \epsilon \left(-\check{p}\mathbf{n} + Pr\frac{\partial \check{\mathbf{u}}}{\partial \mathbf{n}}\right) - \frac{\epsilon}{2}(\check{\mathbf{u}} \cdot \mathbf{n})\check{\mathbf{u}},$$

we can show that $(\check{\mathbf{u}}, \check{\theta}, \check{S}, \check{\mathbf{g}}_\epsilon) \in \mathcal{A}_{\mathrm{ad}}^\epsilon$ and thus

$$\mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}_\epsilon) \ge \mathcal{J}(\mathbf{u}_\epsilon, \mathbf{g}_\epsilon), \tag{3.9}$$

where $(\mathbf{u}_\epsilon, \theta_\epsilon, S_\epsilon, \mathbf{g}_\epsilon)$ is an optimal element of $(P)_\epsilon$. Expanding the left-side of (3.9), we obtain

$$
\begin{aligned}
\mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) \geq {}& \mathcal{J}(\mathbf{u}_\epsilon, \mathbf{g}_\epsilon) - \frac{\epsilon^2 \gamma}{2} \int_0^T \int_\Gamma |-\check{p}\mathbf{n} + Pr\frac{\partial \check{\mathbf{u}}}{\partial \mathbf{n}} - \frac{1}{2}(\check{\mathbf{u}} \cdot \mathbf{n})\check{\mathbf{u}}|^2 \, \mathrm{d}\Gamma \, \mathrm{d}t \\
& - \epsilon\gamma \int_0^T \int_\Gamma \check{\mathbf{g}} \cdot (-\check{p}\mathbf{n} + Pr\frac{\partial \check{\mathbf{u}}}{\partial \mathbf{n}} - \frac{1}{2}(\check{\mathbf{u}} \cdot \mathbf{n})\check{\mathbf{u}}) \, \mathrm{d}\Gamma \, \mathrm{d}t.
\end{aligned}
\tag{3.10}
$$

Putting $\epsilon = \epsilon_k$ in (3.10) and taking $k \to \infty$ yields

$$
\mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) \geq \liminf_{k \to \infty} \mathcal{J}(\mathbf{u}_{\epsilon_k}, \mathbf{g}_{\epsilon_k}) \geq \mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}). \qquad \square
$$

3.2. Optimality of the limit

In this section, we will show that the limit $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S})$ defined in Theorem 3.2 is indeed a solution of the optimal Dirichlet control problem (P). We begin by deriving some continuous dependence results for the solutions of (2.5).

Lemma 3.4. *Let* $\mathbf{g} \in \mathbf{W}_1^\Sigma$ *with* $\boldsymbol{\eta}_g(0) \in \mathbb{H}$. *Then there exist a unique solution* $\boldsymbol{\eta} \in \mathbf{W}_1$ *to the non-homogeneous Stokes problem*

$$
\begin{aligned}
(\boldsymbol{\eta}_t, \mathbf{v}) + Pr(\nabla\boldsymbol{\eta}, \nabla\mathbf{v}) &= 0 \quad \forall \mathbf{v} \in \mathbb{V}, \\
\boldsymbol{\eta}|_\Gamma &= \mathbf{g} \\
\boldsymbol{\eta}(\mathbf{x}, 0) &= 0 \qquad in \quad \Omega.
\end{aligned}
$$

Moreover, there is a constant C_0 *depending only on* Ω *such that*

$$
\|\boldsymbol{\eta}\|_{L^\infty(0,T;\mathbf{H})} + \|\boldsymbol{\eta}\|_{W_1} \leq C_0 \|\mathbf{g}\|_{W_1^\Sigma},
\tag{3.11}
$$

where $\mathbb{H}$ *is the closure of* $\mathbb{V}$ *in* L^2*-norm and* $\boldsymbol{\eta}_g(0)$ *is as defined in Section 2.1.*

Proof. See for *e.g.* [16]. $\qquad \square$

Lemma 3.5. *Let* $\mathbf{g} \in \mathbf{W}_1^\Sigma$ *and* $\mathbf{u}_0 \in \mathbf{H}$ *with* $\mathbf{u}_0 - \mathbf{u}_g(0) \in \mathbb{H}$. *Then there exists a unique solution* $(\mathbf{u}, \theta, S) \in W(0, T)$ *to the problem in* (2.5). *Moreover, there is a constant* $\hat{C}_0$ *such that*

$$
\sup_{t \in [0,T]} \|(\mathbf{u}, \theta, S)\|^2 + \int_0^T \|\nabla(\mathbf{u}, \theta, S)\|^2 \, \mathrm{d}t \leq \hat{C}_0 \left[\|\mathbf{g}\|_{W_1^\Sigma}^2 + \|(f_1, f_2)\|_{L^2(H^{-1}(\Omega))}^2 + \|(\mathbf{u}_0, \theta_0, S_0)\|^2 \right].
$$

Proof. The proof is very similar to the one given for Lemma 2.3. Of course, here we have to use Stokes extension defined in Lemma 3.4 first to reformulate the non-homogeneous problem (2.5) into one with homogeneous boundary conditions. Such a reformulation can be carried out, for *e.g.*, as illustrated in the proof of the next lemma. Therefore the proof is omitted. $\qquad \square$

Lemma 3.6. *Suppose* $(\mathbf{u}_1, \theta_1, S_1, \mathbf{g}_1), (\mathbf{u}_2, \theta_2, S_2, \mathbf{g}_2) \in \mathcal{W}(0, T) \times \mathcal{U}_{\mathrm{ad}}$ *are solutions of* (2.5) *with the initial conditions* $(\mathbf{u}_{01}, \theta_{01}, S_{01})$ *and* $(\mathbf{u}_{02}, \theta_{02}, S_{02})$, *respectively. Let* $\mathbf{e}(t) = (\mathbf{e}_u(t), e_\theta(t), e_S(t))$, *where* $\mathbf{e}_u := \mathbf{u}_1 - \mathbf{u}_2$, $e_\theta := \theta_1 - \theta_2$ *and* $e_S := S_1 - S_2$. *Then*

$$
\sup_{t \in [0,T]} \|\mathbf{e}\|^2 + \int_0^T \|\nabla\mathbf{e}\|^2 \, \mathrm{d}t \leq C \|\mathbf{e}(0)\|^2 + \|\mathbf{g}_1 - \mathbf{g}_2\|_{W_1^\Sigma}^2.
$$

Proof. Let $\boldsymbol{\eta}$ be the Stokes extension defined in Lemma 3.4 such that $\boldsymbol{\eta}|_\Gamma = \mathbf{g}_1 - \mathbf{g}_2$. Defining $\boldsymbol{\zeta} = \mathbf{u}_1 - \mathbf{u}_2 - \boldsymbol{\eta}$, $\chi = \theta_1 - \theta_2$ and $\xi = S_1 - S_2$, we see that $(\boldsymbol{\zeta}, \chi, \xi) \in L^2(0, T; \mathbb{V}) \times [L^2(0, T; H^{-1}(\Omega))]^2$ and $\sigma \in L^2(0, T; L_0^2(\Omega))$ satisfy

$$\begin{cases} \partial_t \boldsymbol{\zeta} - Pr\Delta\boldsymbol{\zeta} + \boldsymbol{\zeta} \cdot \nabla\mathbf{u}_1 + \mathbf{u}_2 \cdot \nabla\boldsymbol{\zeta} + \nabla\sigma + (\boldsymbol{\eta} \cdot \nabla)\mathbf{u}_1 + (\mathbf{u}_2 \cdot \nabla)\boldsymbol{\eta} \\ \qquad = Pr^2 Gr_\theta(\chi + N_r\xi)\mathbf{i}_3 \quad \text{in } \Omega, \\ \qquad \nabla \cdot \boldsymbol{\zeta} = 0 \quad \text{in } \Omega, \\ \partial_t\chi - \Delta\chi + (\mathbf{u}_1 \cdot \nabla)\chi + (\boldsymbol{\zeta} \cdot \nabla)\theta_2 + (\boldsymbol{\eta} \cdot \nabla)\theta_2 = 0 \quad \text{in } \Omega, \\ \partial_t\xi - \dfrac{1}{Le}\Delta\xi + (\mathbf{u}_1 \cdot \nabla)\xi + (\boldsymbol{\zeta} \cdot \nabla)S_2 + (\boldsymbol{\eta} \cdot \nabla)S_2 = \dfrac{\alpha_*}{Le}\Delta\chi \quad \text{in } \Omega, \\ \qquad \boldsymbol{\zeta} = \chi = \xi = 0 \quad \text{on } \Gamma, \end{cases} \tag{3.12}$$

and

$$\boldsymbol{\zeta}(\mathbf{x}, 0) = \mathbf{u}_{01} - \mathbf{u}_{02}, \qquad \chi(\mathbf{x}, 0) = \theta_{01} - \theta_{02} \quad \text{and} \quad \xi(\mathbf{x}, 0) = S_{01} - S_{02} \quad \text{in } \Omega.$$

Taking inner product of $(3.12)_1$, $(3.12)_2$ and $(3.12)_3$ with $\boldsymbol{\zeta} \in \mathbb{V}$, $\chi \in H_0^1(\Omega)$ and $\xi \in H_0^1(\Omega)$, respectively, and using skew-symmetry of the nonlinear terms, we obtain

$$\begin{cases} \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\zeta}\|^2 + \dfrac{5Pr}{6}\|\nabla\boldsymbol{\zeta}\|^2 \leq \dfrac{Pr^2 Gr_\theta}{2}[\|\chi\|^2 + N_r\|\xi\|^2] + Pr^2 Gr_\theta\|\boldsymbol{\zeta}\|^2 \\ \qquad\qquad + C\|\boldsymbol{\zeta}\|^2(\|\nabla\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|_{L^4(\Omega)}^4) + \dfrac{1}{2}\|\nabla\boldsymbol{\eta}\|^2 \\ \qquad\qquad + C\|\boldsymbol{\eta}\|^2\|\mathbf{u}_1\|_{L^4(\Omega)}^4, \\ \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\|\chi\|^2 + \dfrac{1}{2}\|\nabla\chi\|^2 \leq \dfrac{Pr}{6\Lambda}\|\nabla\boldsymbol{\zeta}\|^2 + \dfrac{1}{4\Lambda}\|\nabla\boldsymbol{\eta}\|^2 + C(\|\boldsymbol{\eta}\|^2 + \|\boldsymbol{\zeta}\|^2)\|\theta_2\|_{L^4(\Omega)}^4, \\ \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\|\xi\|^2 + \dfrac{1}{2Le}\|\nabla\xi\|^2 \leq \dfrac{Pr}{6}\|\nabla\boldsymbol{\zeta}\|^2 + \dfrac{1}{4}\|\nabla\boldsymbol{\eta}\|^2 + C\|S_2\|_{L^4(\Omega)}^4((\|\boldsymbol{\eta}\|^2 + \|\boldsymbol{\zeta}\|^2) \\ \qquad\qquad + \dfrac{\alpha_*}{Le}\|\nabla\chi\|\|\nabla\xi\|. \end{cases} \tag{3.13}$$

In deriving (3.13), we estimated the trilinear terms $((\boldsymbol{\eta} \cdot \nabla)\mathbf{u}_1, \boldsymbol{\zeta})$, $((\boldsymbol{\zeta} \cdot \nabla)\mathbf{u}_1, \boldsymbol{\zeta})$ and $(\mathbf{u}_2 \cdot \nabla\boldsymbol{\eta}, \boldsymbol{\zeta})$ using (A-1), (A-4) and (A-5). Moreover, the terms $(\boldsymbol{\zeta} \cdot \nabla\theta_2, \xi)$, $(\boldsymbol{\eta} \cdot \nabla\theta_2, \xi)$, $(\boldsymbol{\zeta} \cdot \nabla S_2, \chi)$ and $(\boldsymbol{\eta} \cdot \nabla S_2, \chi)$ were estimated using (A-1).

Multiplying $(3.13)_2$ by Λ and adding it to $(3.13)_3$, and choosing λ_1 and λ_2 as we did in the proof of Lemma 2.3 yields

$$\dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}(\Lambda\|\chi\|^2 + \|\xi\|^2) + \dfrac{\lambda_1}{2}\|\nabla\chi\|^2 + \dfrac{\lambda_2}{2}\|\nabla\xi\|^2 \leq \dfrac{Pr}{3}\|\nabla\boldsymbol{\zeta}\|^2 + \dfrac{1}{2}\|\nabla\boldsymbol{\eta}\|^2 \\ \qquad\qquad + C(\|\boldsymbol{\eta}\|^2 + \|\boldsymbol{\zeta}\|^2)\left(\|\theta_2\|_{L^4(\Omega)}^4 + \|S_2\|_{L^4(\Omega)}^4\right). \tag{3.14}$$

Adding (3.14) and $(3.13)_1$, we obtain

$$\dfrac{\mathrm{d}}{\mathrm{d}t}\|(\boldsymbol{\zeta}, \Lambda\chi, \xi)\|^2 + Pr\|\nabla\boldsymbol{\zeta}\|^2 + \lambda_1\|\nabla\chi\|^2 + \lambda_2\|\nabla\xi\|^2 \leq G_1(t)\|(\boldsymbol{\zeta}, \Lambda\chi, \xi)\|^2 + G_2(t), \tag{3.15}$$

where

$$G_1(t) := C_1\left[1 + \|\nabla\mathbf{u}_1\|^2 + \|(\mathbf{u}_2, \theta_2, S_2)\|_{L^4(\Omega)}^4 + \|\mathbf{u}_1\|_{L^4(\Omega)}^4\right]$$
$$G_2(t) := C_2\left[\|\nabla\boldsymbol{\eta}\|^2 + \|\boldsymbol{\eta}\|^2\left(\|\mathbf{u}_1\|_{L^4(\Omega)}^4 + \|\theta_2\|_{L^4(\Omega)}^4 + \|S_2\|_{L^4(\Omega)}^4\right)\right].$$

Notice that by Lemmas 3.4–3.5 and (2.1), $G_i(t) \in L^1(0, T)$, $i = 1, 2$. Therefore desired estimate now follows by Gronwall lemma applied to (3.15). □

Theorem 3.7. *Let $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S})$ be the limit defined in Theorem 3.2. Then $(\hat{\mathbf{u}}, \hat{\theta}, \hat{S})$ is a solution of the optimal Dirichlet control problem (P).*

Proof. Let $(\check{\mathbf{u}}, \check{\theta}, \check{S}, \check{\mathbf{g}}) \in \mathcal{A}_{\mathrm{ad}}$ be an arbitrary element. We would like to show $\mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) \geq \mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}})$. Using the denseness of the space $C^\infty((0,T) \times \Gamma) \cap \mathbf{W}_1^\Sigma$ in $\mathbf{W}_1^\Sigma$, we may choose a sequence $\{\check{\mathbf{g}}_i\}_{i=1}^\infty \subset C^\infty((0,T) \times \Gamma) \cap \mathbf{W}_1^\Sigma$ such that

$$\|\check{\mathbf{g}}_i - \check{\mathbf{g}}\|_{W_1^\Sigma} \to 0 \quad \text{as} \quad i \to \infty. \tag{3.16}$$

Similarly, using the denseness of the space $C^\infty(\Omega) \cap [\mathbf{H} \times (L^2(\Omega))^2]$ in $\mathbf{H} \times (L^2(\Omega))^2$, we may choose a sequence $\{(\mathbf{u}_{0i}, \theta_{0i}, S_{0i})\}_{i=1}^\infty \subset C^\infty(\Omega) \cap [\mathbf{H} \times (L^2(\Omega))^2]$ such that

$$\|(\mathbf{u}_{0i}, \theta_{0i}, S_{0i}) - (\mathbf{u}_0, \theta_0, S_0)\| \to 0 \quad \text{as} \quad i \to \infty. \tag{3.17}$$

Now let $(\check{\mathbf{u}}_i, \check{\theta}_i, \check{S}_i)$ be the solution to (2.5) with boundary conditions $\check{\mathbf{u}}_i|_\Gamma = \check{\mathbf{g}}_i$ $\check{\theta}_i|_\Gamma = 0$ and $\check{S}_i|_\Gamma = 0$, and initial conditions $\check{\mathbf{u}}_i(\mathbf{x}, 0) = \mathbf{u}_{0i}$, $\check{\theta}_i(\mathbf{x}, 0) = \theta_{0i}$ and $\check{S}_i(\mathbf{x}, 0) = S_{0i}$. Using Lemma 3.6, (3.16) and (3.17), we have that

$$\|(\check{\mathbf{u}}_i, \check{\theta}_i, \check{S}_i) - (\check{\mathbf{u}}, \check{\theta}, \check{S})\|_{W(0,T)} \to 0 \quad \text{as} \quad i \to \infty.$$

The regularity of solutions yields that $(\check{\mathbf{u}}_i, \check{\theta}_i, \check{S}_i) \in L^2(\mathbf{H}^{3/2}(\Omega)) \times (L^2(H^{3/2}(\Omega)))^2$ and pressure $\check{p}_i \in L^2(H^{1/2}(\Omega) \cap L_0^2(\Omega))$. Thus

$$-\check{p}_i\mathbf{n} + Pr\frac{\partial \check{\mathbf{u}}_i}{\partial \mathbf{n}} - \frac{1}{2}(\check{\mathbf{u}}_i \cdot \mathbf{n})\check{\mathbf{u}}_i \in L^2(\mathbf{L}^2(\Gamma)).$$

By Lemma 3.3 we obtain $\mathcal{J}(\check{\mathbf{u}}_i, \check{\mathbf{g}}_i) \geq \mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}})$. Taking limit as $i \to \infty$, we obtain $\mathcal{J}(\check{\mathbf{u}}, \check{\mathbf{g}}) = \lim_{i \to \infty} \mathcal{J}(\check{\mathbf{u}}_i, \check{\mathbf{g}}_i) \geq \mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}})$. □

4. First order necessary conditions of optimality

In this section, we will characterize the optimal controls of $(P)_\epsilon$ by giving the necessary conditions of optimality. We begin by verifying the Gâteaux differentiability of the state $(\mathbf{u}, \theta, S)$ with respect to the control $\mathbf{g}$. The Gâteaux derivative plays an important role in the characterization of optimality condition and are useful in determining the sensitivity of the system at a particular point $(\mathbf{u}, \theta, S)$ in a particular direction $\mathbf{h}$. First we prove the control-to-state mapping from $\mathcal{U}_{\mathrm{ad}}^\epsilon$ to $\mathcal{W}(0,T)$ is continuous.

Lemma 4.1. *Assume $\epsilon \in (0, \frac{1}{Pr})$. The control-to-state mapping from $\mathcal{U}_{\mathrm{ad}}^\epsilon$ to $\mathcal{W}(0,T)$, defined by $\mathbf{g} \mapsto (\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g}))$, where $(\mathbf{u}, \theta, S)$ is the solution to (2.7), satisfy*

$$\lim_{\lambda \to 0} \|(\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}), \theta(\mathbf{g} + \lambda\mathbf{h}), S(\mathbf{g} + \lambda\mathbf{h})) - (\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g}))\|_{W(0,T)} \to 0, \tag{4.1}$$

where $\mathbf{h} \in \mathcal{U}_{\mathrm{ad}}^\epsilon$ is arbitrarily fixed.

Proof. Setting $(\tilde{\mathbf{u}}, \tilde{\theta}, \tilde{S}) := (\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}), \theta(\mathbf{g} + \lambda\mathbf{h}), S(\mathbf{g} + \lambda\mathbf{h})) - (\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g}))$, notice that $(\tilde{\mathbf{u}}, \tilde{\theta}, \tilde{S})$ satisfies the system

$$\begin{cases} (\partial_t\tilde{\mathbf{u}}, \mathbf{v}) + Pr(\nabla\tilde{\mathbf{u}}, \nabla\mathbf{v}) + c(\tilde{\mathbf{u}}, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}), \tilde{\mathbf{u}}, \mathbf{v}) + \frac{1}{\epsilon}(\tilde{\mathbf{u}}, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\tilde{\theta} + N_r\tilde{S})\mathbf{i}_3, \mathbf{v}) + \frac{\lambda}{\epsilon}(\mathbf{h}, \mathbf{v})_\Gamma, \\[2mm] (\partial_t\tilde{\theta}, \phi) + (\nabla\tilde{\theta}, \nabla\phi) + (\tilde{\mathbf{u}} \cdot \nabla\theta, \phi) + (\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}) \cdot \nabla\tilde{\theta}, \phi) = 0, \\[2mm] (\partial_t\tilde{S}, \psi) + \frac{1}{Le}(\nabla\tilde{S}, \nabla\psi) + (\tilde{\mathbf{u}} \cdot \nabla S, \psi) + (\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}) \cdot \nabla\tilde{S}, \psi) = \frac{\alpha_*}{Le}(\nabla\tilde{\theta}, \nabla\psi), \end{cases} \tag{4.2}$$

and

$$\tilde{\mathbf{u}}(\mathbf{x}, 0) = 0, \quad \tilde{\theta}(\mathbf{x}, 0) = 0 \quad \text{and} \quad \tilde{S}(\mathbf{x}, 0) = 0,$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$. Therefore we can use energy arguments as in the uniqueness part of the proof of Lemma 2.3 and obtain

$$
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|(\tilde{\mathbf{u}}, \Lambda\tilde{\theta}, \tilde{S})\|^2 + \frac{Pr}{4}\|\nabla\tilde{\mathbf{u}}\|^2 + \lambda_1\|\nabla\tilde{\theta}\|^2 + \lambda_2\|\nabla\tilde{S}\|^2 + \frac{1}{2\epsilon}\|\tilde{\mathbf{u}}\|_{0,\Gamma}^2
$$
$$
\leq \|(\tilde{\mathbf{u}}, \Lambda\tilde{\theta}, \tilde{S})\|^2 G_3(t) + \frac{\lambda^2}{\epsilon}\|\mathbf{h}\|_{0,\Gamma}^2, \tag{4.3}
$$

where $G_3(t) := C[1 + \|(\mathbf{u}, \theta, S)\|_{L^4(\Omega)}^4 + \|\nabla\mathbf{u}\|^2]$ and $G_3(t) \in L^1(0,T)$, by (2.1) and Lemma 2.3. Thus by Gronwall's inequality, we obtain

$$
\|(\tilde{\mathbf{u}}, \Lambda\tilde{\theta}, \tilde{S})\|_{L^\infty(\mathbf{H}\times(L^2(\Omega))^2)\cap L^2(\mathbf{V}\times(H_0^1(\Omega))^2)} \leq \frac{C\lambda}{\sqrt{\epsilon}}\|\mathbf{h}\|_{L^2(\mathbf{L}^2(\Gamma))}. \tag{4.4}
$$

Moreover, we can show from (4.2) using the estimates (4.4) and supremum arguments that

$$
\|(\partial_t\tilde{\mathbf{u}}, \partial_t\tilde{\theta}, \partial_t\tilde{S})\|_{L^2(\mathbb{V}^*\times(H^{-1}(\Omega))^2)} \leq C\frac{\lambda}{\sqrt{\epsilon}}\|\mathbf{h}\|_{L^2(\mathbf{L}^2(\Gamma))}. \qquad \square
$$

Lemma 4.2. *Assume $\epsilon \in (0, \frac{1}{Pr})$. The control-to-state mapping from $\mathcal{U}_{\mathrm{ad}}^\epsilon$ to $\mathcal{W}(0,T)$, defined by $\mathbf{g} \mapsto (\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g}))$, where $(\mathbf{u}, \theta, S)$ is the solution to (2.7), has a Gâteaux derivative $D(\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g})) \cdot \mathbf{h}$ in every direction $\mathbf{h} \in \mathcal{U}_{\mathrm{ad}}^\epsilon$. Moreover, $(\mathbf{u}', \theta', S') = D(\mathbf{u}(\mathbf{g}), \theta(\mathbf{g}), S(\mathbf{g})) \cdot \mathbf{h}$ is the weak solution of the following problem.*

$$
\begin{cases}
(\partial_t\mathbf{u}', \mathbf{v}) + Pr(\nabla\mathbf{u}', \nabla\mathbf{v}) + c(\mathbf{u}', \mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{u}', \mathbf{v}) + \dfrac{1}{\epsilon}(\mathbf{u}', \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\theta' + N_r S')\mathbf{i}_3, \mathbf{v}) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \dfrac{1}{\epsilon}(\mathbf{h}, \mathbf{v})_\Gamma, \\
(\partial_t\theta', \phi) + (\nabla\theta', \nabla\phi) + (\mathbf{u}' \cdot \nabla\theta, \phi) + (\mathbf{u} \cdot \nabla\theta', \phi) = 0, \\
(\partial_t S', \psi) + \dfrac{1}{Le}(\nabla S', \nabla\psi) + (\mathbf{u}' \cdot \nabla S, \psi) + (\mathbf{u} \cdot \nabla S', \psi) = \dfrac{\alpha_*}{Le}(\nabla\theta', \nabla\psi),
\end{cases} \tag{4.5}
$$

and

$$
\mathbf{u}'(\mathbf{x}, 0) = 0, \quad \theta'(\mathbf{x}, 0) = 0 \quad and \quad S'(\mathbf{x}, 0) = 0,
$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$.

Proof. Let $\lambda \in (-1, 1)$ and $\lambda \neq 0$. For $(\tilde{\mathbf{u}}, \tilde{\theta}, \tilde{S})$ defined as in Lemma 4.1, define $(\mathbf{u}_\lambda, \theta_\lambda, S_\lambda) = (\tilde{\mathbf{u}}/\lambda, \tilde{\theta}/\lambda, \tilde{S}/\lambda)$. Then we have that

$$
\begin{cases}
(\partial_t\mathbf{u}_\lambda, \mathbf{v}) + Pr(\nabla\mathbf{u}_\lambda, \nabla\mathbf{v}) + c(\mathbf{u}_\lambda, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}), \mathbf{u}_\lambda, \mathbf{v}) + \dfrac{1}{\epsilon}(\mathbf{u}_\lambda, \mathbf{v})_\Gamma \\
\qquad\qquad\qquad\qquad = (Pr^2 Gr_\theta(\theta_\lambda + N_r S_\lambda)\mathbf{i}_3, \mathbf{v}) + \dfrac{1}{\epsilon}(\mathbf{h}, \mathbf{v})_\Gamma, \\
(\partial_t\theta_\lambda, \phi) + (\nabla\theta_\lambda, \nabla\phi) + (\mathbf{u}_\lambda \cdot \nabla\theta, \phi) + (\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}) \cdot \nabla\theta_\lambda, \phi) = 0, \\
(\partial_t S_\lambda, \psi) + \dfrac{1}{Le}(\nabla S_\lambda, \nabla\psi) + (\mathbf{u}_\lambda \cdot \nabla S, \psi) + (\mathbf{u}(\mathbf{g} + \lambda\mathbf{h}) \cdot \nabla S_\lambda, \psi) = \dfrac{\alpha_*}{Le}(\nabla\theta_\lambda, \nabla\psi),
\end{cases} \tag{4.6}
$$

and

$$
\mathbf{u}_\lambda(\mathbf{x}, 0) = 0, \quad \theta_\lambda(\mathbf{x}, 0) = 0 \quad and \quad S_\lambda(\mathbf{x}, 0) = 0,
$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$. It can now be shown using energy arguments as in the proof of Lemma 4.1 that

$$
\|(\mathbf{u}_\lambda, \Lambda\theta_\lambda, S_\lambda)\|_{L^\infty(\mathbf{H}\times(L^2(\Omega))^2)\cap L^2(\mathbf{V}\times(H_0^1(\Omega))^2)} + \frac{1}{\epsilon}\|\mathbf{u}_\lambda\|_{L^2(\mathbf{L}^2(\Gamma))} \leq \frac{C}{\sqrt{\epsilon}}\|\mathbf{h}\|_{L^2(\mathbf{L}^2(\Gamma))} \tag{4.7}
$$

and

$$\|(\partial_t \mathbf{u}_\lambda, \partial_t \theta_\lambda, \partial_t S_\lambda)\|_{L^2(\mathbb{V}^* \times (H^{-1}(\Omega))^2)} \le C \frac{1}{\sqrt{\epsilon}} \|\mathbf{h}\|_{L^2(\mathbf{L}^2(\Gamma))}.$$

Therefore we can extract a sequence $\{(\mathbf{u}_{\lambda_k}, \theta_{\lambda_k}, S_{\lambda_k})\}_{k=1}^\infty$ and $\{\lambda_k\}_{k=1}^\infty \subset (-1, 1)$ such that as $k \to \infty$,

$$\lambda_k \to 0, \quad \mathbf{u}_{\lambda_k} \rightharpoonup \mathbf{u}' \text{ in } L^2(\mathbf{V}), \quad \theta_{\lambda_k} \rightharpoonup \theta' \text{ in } L^2(H_0^1(\Omega)), \quad S_{\lambda_k} \rightharpoonup S' \text{ in } L^2(H_0^1(\Omega))$$

and $\mathbf{u}_n|_\Gamma \rightharpoonup \mathbf{u}'|_\Gamma$ in $L^2(L^2(\Gamma))$. Moreover, by Aubin–Simon compactness Lemma ([27], Cor. 4, p. 85),

$$\mathbf{u}_{\lambda_k} \to \mathbf{u}' \text{ in } L^2(\mathbf{H}), \quad \theta_{\lambda_k} \to \theta' \text{ in } L^2(L^2(\Omega)), \quad S_{\lambda_k} \to S' \text{ in } L^2(L^2(\Omega)).$$

Setting $\lambda = \lambda_k$ in (4.6) and letting $k \to \infty$, we can show $(\mathbf{u}', \theta', S')$ indeed satisfy (4.5) by arguing as in the proof of Lemma 4.1.

Let us next prove the strong convergence of $(\mathbf{u}_\lambda, \theta_\lambda, S_\lambda)$ also in the topology of $W(0, T)$. Subtracting (4.5) from (4.6) and denoting $(\mathbf{u}_\lambda, \theta_\lambda, S_\lambda) - (\mathbf{u}', \theta', S')$ by $(\bar{\mathbf{u}}, \bar{\theta}, \bar{S})$ satisfy

$$\begin{cases} (\partial_t \bar{\mathbf{u}}, \mathbf{v}) + Pr(\nabla \bar{\mathbf{u}}, \nabla \mathbf{v}) + c(\tilde{\mathbf{u}}, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \bar{\mathbf{u}}, \mathbf{v}) + c(\tilde{\mathbf{u}}, \mathbf{u}_\lambda, \mathbf{v}) + \frac{1}{\epsilon}(\bar{\mathbf{u}}, \mathbf{v})_\Gamma = (Pr^2 Gr_\theta(\bar{\theta} + N_r \bar{S})\mathbf{i}_3, \mathbf{v}), \\ (\partial_t \bar{\theta}, \phi) + (\nabla \bar{\theta}, \nabla \phi) + (\bar{\mathbf{u}} \cdot \nabla \theta, \phi) + (\mathbf{u} \cdot \nabla \bar{\theta}, \phi) + (\tilde{\mathbf{u}} \cdot \nabla \theta_\lambda, \phi) = 0, \\ (\partial_t \bar{S}, \psi) + \frac{1}{Le}(\nabla \bar{S}, \nabla \psi) + (\bar{\mathbf{u}} \cdot \nabla S, \psi) + (\mathbf{u} \cdot \nabla \bar{S}, \psi) + (\tilde{\mathbf{u}} \cdot \nabla S_\lambda, \psi) = \frac{\alpha_*}{Le}(\nabla \bar{\theta}, \nabla \psi), \end{cases}$$

(4.8)

and

$$\bar{\mathbf{u}}(\mathbf{x}, 0) = 0, \quad \bar{\theta}(\mathbf{x}, 0) = 0 \quad \text{and} \quad \bar{S}(\mathbf{x}, 0) = 0,$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$. We will again use energy methods. By setting $\mathbf{v} = \bar{\mathbf{u}}$ in $(4.8)_1$, we can estimate the trilinear terms $c(\bar{\mathbf{u}}, \mathbf{u}, \bar{\mathbf{u}})$ and $c(\tilde{\mathbf{u}}, \mathbf{u}_\lambda, \bar{\mathbf{u}})$ using (A-2)$-$(A-3). Also, setting $\phi = \bar{\theta}$ and $\psi = \bar{S}$ in $(4.8)_2$ and $(4.8)_3$, respectively, we can estimate the trilinear terms using (A-1). Therefore, we obtain

$$\begin{cases} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{\mathbf{u}}\|^2 + \frac{3Pr}{4}\|\nabla \bar{\mathbf{u}}\|^2 + \frac{1}{\epsilon}\|\bar{\mathbf{u}}\|_{0,\Gamma}^2 \le C\|(\bar{\mathbf{u}}, \Lambda\bar{\theta}, \bar{S})\|^2 + C\|\bar{\mathbf{u}}\|^2[\|\nabla \mathbf{u}\|^2 + \|\nabla \mathbf{u}_\lambda\|^2 + \|\mathbf{u}\|_{L^4(\Omega)}^4] \\ \qquad\qquad\qquad\qquad + C\|\tilde{\mathbf{u}}\|_{L^4(\Omega)}^2(\|\nabla \mathbf{u}_\lambda\| + \|\mathbf{u}_\lambda\|_{L^4(\Omega)}^2) \\ \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{\theta}\|^2 + \frac{1}{2}\|\nabla \bar{\theta}\|^2 \le \frac{Pr}{8\Lambda}\|\nabla \bar{\mathbf{u}}\|^2 + C[\|\bar{\mathbf{u}}\|^2\|\theta\|_{L^4(\Omega)}^4 + \|\tilde{\mathbf{u}}\|^2\|\theta_\lambda\|_{L^4(\Omega)}^4] + \frac{1}{2}\|\nabla \tilde{\mathbf{u}}\|^2 \\ \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{S}\|^2 + \frac{1}{2Le}\|\nabla \bar{S}\|^2 \le \frac{Pr}{8}\|\nabla \bar{\mathbf{u}}\|^2 + C[\|\bar{\mathbf{u}}\|^2\|S\|_{L^4(\Omega)}^4 + \|\tilde{\mathbf{u}}\|^2\|S_\lambda\|_{L^4(\Omega)}^4] + \frac{1}{2}\|\nabla \tilde{\mathbf{u}}\|^2 \\ \qquad\qquad\qquad\qquad + \frac{\alpha_*}{Le}\|\nabla \bar{\theta}\|\|\nabla \bar{S}\|. \end{cases}$$

(4.9)

Multiplying $(4.9)_2$ by Λ and adding the results to $(4.9)_3$ yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\Lambda\|\bar{\theta}\|^2 + \|\bar{S}\|^2) + \lambda_1\|\nabla \bar{\theta}\|^2 + \lambda_2\|\nabla \bar{S}\|^2 \le \frac{Pr}{4}\|\nabla \bar{\mathbf{u}}\|^2 + C(\|\bar{\mathbf{u}}\|^2 + \|\tilde{\mathbf{u}}\|^2)(\|\theta\|_{L^4(\Omega)}^4 + \|S\|_{L^4(\Omega)}^4) \\ + C\|\nabla \tilde{\mathbf{u}}\|^2.$$

(4.10)

Combining $(4.9)_1$ and (4.10), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|(\bar{\mathbf{u}}, \Lambda\bar{\theta}, \bar{S})\|^2 + \lambda_1\|\nabla \bar{\theta}\|^2 + \lambda_2\|\nabla \bar{S}\|^2 + \frac{Pr}{2}\|\nabla \bar{\mathbf{u}}\|^2 + \frac{1}{\epsilon}\|\bar{\mathbf{u}}\|_{0,\Gamma}^2 \le G_4(t)\|(\bar{\mathbf{u}}, \Lambda\bar{\theta}, \bar{S})\|^2 + G_5(t),$$

where

$$G_4(t) := C\left[1 + \|\nabla \mathbf{u}\|^2 + \|\nabla \mathbf{u}_\lambda\|^2 + \|(\mathbf{u}, \theta, S)\|_{L^4(\Omega)}^4\right]$$

$$G_5(t) := C\left[\|\nabla \tilde{\mathbf{u}}\|^2(1 + \|\mathbf{u}_\lambda\|^2) + \|\tilde{\mathbf{u}}\|^2\left(\|\nabla \mathbf{u}_\lambda\|^2 + \|(\theta, S)\|_{L^4(\Omega)}^4\right)\right].$$

Since $G_4(t), G_5(t) \in L^1(0, T)$, we can use Gronwall lemma and obtain

$$\|(\bar{\mathbf{u}}, \Lambda\bar{\theta}, \bar{S})\| \leq C\frac{\lambda}{\sqrt{\epsilon}}\|\mathbf{h}\|_{L^2(L^2(\Gamma))}. \tag{4.11}$$

Moreover, we can show from (4.8) using the estimates (4.4), (4.7) and (4.11), and supremum argument that

$$\|(\partial_t\bar{\mathbf{u}}, \partial_t\bar{\theta}, \partial_t\bar{S})\|_{L^2(\mathbb{V}^* \times (H^{-1}(\Omega))^2)} \leq C\frac{\lambda}{\sqrt{\epsilon}}\|\mathbf{h}\|_{L^2(\mathbf{L}^2(\Gamma))}.$$

Therefore the desired strong convergence now follows, *i.e.*,

$$\left\|\frac{1}{\lambda}[(\mathbf{u}_{g+\lambda h}, \theta_{g+\lambda h}, S_{g+\lambda h}) - (\mathbf{u}_g, \theta_g, S_g) - \lambda(\mathbf{u}', \theta', S')]\right\|_{\mathcal{W}(0,T)} \to 0 \quad \text{as} \quad \lambda \to 0. \qquad \square$$

Recall that if

$$\mathcal{J}(\mathbf{u}^*(\mathbf{g}^*), \mathbf{g}^*) = \inf_{\mathbf{g}\in\mathcal{U}_{\mathrm{ad}}^\epsilon} \mathcal{J}(\mathbf{u}(\mathbf{g}), \mathbf{g})$$

and the cost function is Gâteaux differentiable, then necessarily

$$DJ(\mathbf{u}^*(\mathbf{g}^*), \mathbf{g}^*)\mathbf{h} = 0 \quad \text{for all } \mathbf{h} \in \mathcal{U}_{\mathrm{ad}}^\epsilon, \tag{4.12}$$

where $DJ(\mathbf{u}^*(\mathbf{g}^*), \mathbf{g}^*)$ denotes the Gâteaux derivative of $\mathcal{J}$ at $\mathbf{g}^* \in \mathcal{U}_{\mathrm{ad}}^\epsilon$. From (4.12), we have

$$\int_0^T (\Theta_u, \mathbf{u}')\,\mathrm{d}t + \gamma\int_0^T \int_\Gamma \mathbf{gh}\,\mathrm{d}\Gamma\,\mathrm{d}t = 0 \quad \forall\mathbf{h} \in \mathcal{U}_{\mathrm{ad}}^\epsilon. \tag{4.13}$$

In practical computation of optimal control, gradient of the cost functional should be computed without the need to evaluate the state variables. This can be achieved by employing adjoint variables instead of sensitivities.

Theorem 4.3. *Assume $\epsilon \in (0, \frac{1}{Pr})$. If $(\mathbf{u}, \theta, S, \mathbf{g})$ be an optimal solution for $(P)_\epsilon$, then we have*

$$\frac{1}{\epsilon}\int_0^T \int_\Gamma \breve{\mathbf{u}}\mathbf{h}\,\mathrm{d}\Gamma\,\mathrm{d}t + \gamma\int_0^T \int_\Gamma \mathbf{gh}\,\mathrm{d}\Gamma\,\mathrm{d}t = 0 \quad \forall\mathbf{h} \in \mathcal{U}_{\mathrm{ad}}^\epsilon, \tag{4.14}$$

where $(\breve{\mathbf{u}}, \breve{\theta}, \breve{S})$ is the weak solution of the adjoint equations

$$\begin{cases} -(\partial_t\breve{\mathbf{u}}, \mathbf{v}) + Pr(\nabla\breve{\mathbf{u}}, \nabla\mathbf{v}) + c(\mathbf{u}, \mathbf{v}, \breve{\mathbf{u}}) + c(\mathbf{v}, \mathbf{u}, \breve{\mathbf{u}}) + \frac{1}{\epsilon}(\breve{\mathbf{u}}, \mathbf{v})_\Gamma \\ \qquad\qquad + (\mathbf{v} \cdot \nabla S, \breve{S}) + (\mathbf{v} \cdot \nabla\theta, \breve{\theta}) = (\Theta_u(\mathbf{u}), \mathbf{v}) \\ -(\partial_t\breve{\theta}, \phi) + (\nabla\breve{\theta}, \nabla\phi) + (\mathbf{u} \cdot \nabla\phi, \breve{\theta}) - Pr^2Gr_\theta(\breve{\mathbf{u}}, \mathbf{i}_3\phi) = \frac{\alpha_*}{Le}(\nabla\breve{S}, \nabla\phi) \\ -(\partial_t\breve{S}, \psi) + \frac{1}{Le}(\nabla\breve{S}, \nabla\psi) + (\mathbf{u} \cdot \nabla\psi, \breve{S}) - Pr^2Gr_\theta N_r(\breve{\mathbf{u}}, \mathbf{i}_3\psi) = 0, \end{cases} \tag{4.15}$$

for all $(\mathbf{v}, \phi, \psi) \in \mathbf{V} \times (H_0^1(\Omega))^2$, and $\breve{\mathbf{u}}(\mathbf{x}, T) = \breve{\theta}(\mathbf{x}, T) = \breve{S}(\mathbf{x}, T) = 0$.

Proof. Let $(\mathbf{u}, \theta, S, \mathbf{g})$ be an optimal solution of the control problem $(P)_\epsilon$ and let $\mathbf{u}'$ be as defined by (4.5). Then setting $\mathbf{v} = \mathbf{u}'$ in $(4.15)_1$ and using the optimality condition (4.14) yields

$$\begin{aligned} 0 = \gamma\int_0^T \int_\Gamma \mathbf{g}_\epsilon\mathbf{h}\,\mathrm{d}\Gamma &- (\partial_t\breve{\mathbf{u}}, \mathbf{u}') + Pr(\nabla\breve{\mathbf{u}}, \nabla\mathbf{u}') + c(\mathbf{u}, \mathbf{u}', \breve{\mathbf{u}}) + c(\mathbf{u}', \mathbf{u}, \breve{\mathbf{u}}) \\ &+ \frac{1}{\epsilon}(\breve{\mathbf{u}}, \mathbf{u}')_\Gamma + (\mathbf{u}' \cdot \nabla S, \breve{S}) + (\mathbf{u}' \cdot \nabla\theta, \breve{\theta})\,\mathrm{d}t. \end{aligned} \tag{4.16}$$

Setting $\phi = \breve{\theta}$ and $\psi = \breve{S}$ in $(4.5)_2$ and $(4.5)_3$, respectively, and integrating by parts on the time derivative term yield

$$\begin{cases} \int_0^T (\mathbf{u}' \cdot \nabla\theta, \breve{\theta})\, \mathrm{d}t = \int_0^T (\partial_t \breve{\theta}, \theta') - (\nabla\theta', \nabla\breve{\theta}) - (\mathbf{u} \cdot \nabla\theta', \breve{\theta})\, \mathrm{d}t \\ \int_0^T (\mathbf{u}' \cdot \nabla S, \breve{S})\, \mathrm{d}t = \int_0^T (\partial_t \breve{S}, S') - \frac{1}{Le}(\nabla S', \nabla\breve{S}) - (\mathbf{u} \cdot \nabla S', \breve{S}) + \frac{\alpha_*}{Le}(\nabla\theta', \nabla\breve{S})\, \mathrm{d}t. \end{cases} \quad (4.17)$$

On the other hand by setting $\phi = \theta'$ and $\psi = S'$ in $(4.15)_2$ and $(4.15)_3$, respectively, we have

$$\begin{cases} \int_0^T (\partial_t \breve{\theta}, \theta') - (\nabla\breve{\theta}, \nabla\theta') - (\mathbf{u} \cdot \nabla\theta', \breve{\theta})\, \mathrm{d}t = -\int_0^T Pr^2 Gr_\theta(\breve{\mathbf{u}}, \theta') - \frac{\alpha_*}{Le}(\nabla\breve{S}, \nabla\theta')\, \mathrm{d}t \\ \int_0^T (\partial_t \breve{S}, S') - \frac{1}{Le}(\nabla\breve{S}, \nabla S') - (\mathbf{u} \cdot S', \breve{S})\, \mathrm{d}t = -Pr^2 Gr_\theta N_r \int_0^T (\breve{\mathbf{u}}, S')\, \mathrm{d}t. \end{cases} \quad (4.18)$$

Employing (4.18) in (4.17), we obtain

$$\int_0^T (\mathbf{u}' \cdot \nabla\theta, \breve{\theta}) + (\mathbf{u}' \cdot \nabla S, \breve{S})\, \mathrm{d}t = -Pr^2 Gr_\theta \int_0^T (\breve{\mathbf{u}}, \theta' + N_r S')\, \mathrm{d}t \quad (4.19)$$

Therefore by employing (4.19) in (4.16) and integrating by parts on the time derivative term, and using $(4.5)_1$, we obtain the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5. Numerical results

In fluids having two constituents with different molecular diffusivities, the phenomenon of doubly diffusive convection occurs provided the Grashof number is sufficiently large. A prototype model for doubly diffusive convection is a flow in a rectangular enclosure driven by either horizontal or vertical temperature and concentration gradients. Convection that appears in the form of circulations can have undesirable effects in some systems. For example during crystal growth, it affects the local growth conditions and transport rate [5]. Recirculating flows can cause non-uniformities and make reactants from previous cycle to remain in the recirculating region and become the undesired impurities.

We consider a doubly diffusive flow in a square cavity where horizontal temperature and concentration gradients contribute to the fluid buoyancy in opposite ways. Initially fluid in the cavity is at rest at uniform temperature and concentration ($\theta = S = 0$). The no slip boundary condition is assumed on the cavity walls. Two different thermal and concentration boundary conditions are used at the cavity walls. On the left and right walls $\theta = S = 1$ and $\theta = S = 0$ are applied, respectively. The top and bottom walls are assumed to be adiabatic with zero heat and concentration flux. Notice that this configuration corresponds to unstable motion free state ($\mathbf{u} = 0$ and $\theta = S = 1 - x$) for Grashof number Gr_θ less than a critical Grashof number Gr_θ^c.

In this numerical example, we set the Grashof number to $Gr_\theta = 10^4$, Prandtl number $Pr = 1.0$, Lewis number $Le = 2$ and buoyancy ratio $N_r = -1$, which corresponds to supercritical state for which convective flow will occur in the uncontrolled case, see Figure 1. It shows the uncontrolled velocity, temperature and concentration at $t = 2$. In the beginning convective flow develops and after a short transition period reaches a steady state at about time $t = 2$. The simulations have been performed by discretizing the governing equations by mixed Galerkin finite element for spatial discretization on a triangular grid with 1800 elements. The velocity, temperature and concentration are approximated using piecewise quadratic polynomials while the pressure is approximated using piecewise linear polynomials. 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) [25].

The goal of the control here is to keep the system at the unstable motion free state for supercritical Grashof numbers. We formulate this problem as an optimal control problem where the deviation of the state from the no motion state is minimized in a finite time horizon. The control action $\mathbf{g}$ is taken to be the tangential boundary
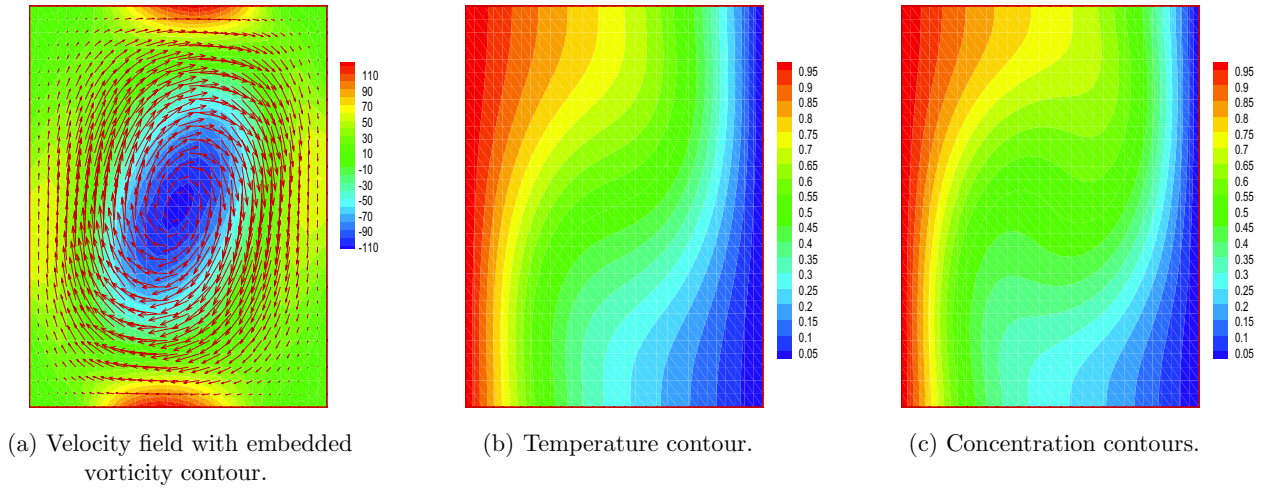
(a) Velocity field with embedded
vorticity contour.

(b) Temperature contour.

(c) Concentration contours.

FIGURE 1. Uncontrolled velocity, temperature and concentration at $t = 2$.

(a) Velocity field with embedded
vorticity contour.
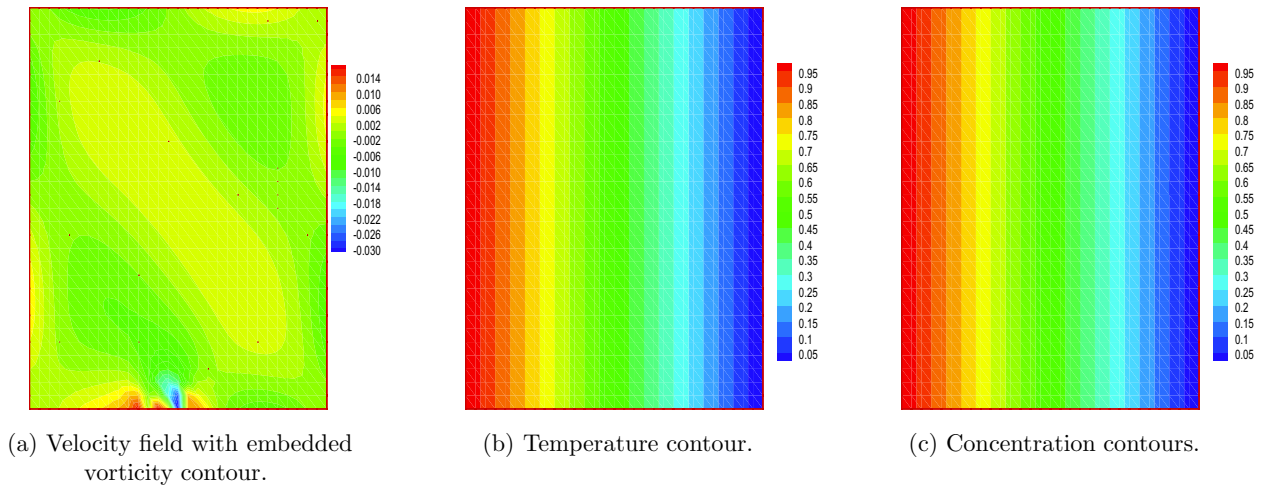
(b) Temperature contour.

(c) Concentration contours.

FIGURE 2. Controlled velocity, temperature and concentration at $t = 2$.

velocity on a small slot (one-tenth of the bottom wall length) on the bottom wall of the cavity. The cost function was defined as in (3.1) with $\Theta(\mathbf{u}) = \frac{\delta^*}{2}\|\nabla\mathbf{u}\|^2 + \delta\|\mathbf{u}\|^2$ and $\mathbf{g} = (g, 0)$, where δ^* is taken to be negligibly small (*e.g.* 10^{-5}) as our interest here really is to minimize the latter term. In fact, our numerical experiments show δ^* can even be taken to be zero for computational purposes. The control that minimizes the cost functional was found using a variable step gradient algorithm, where gradient of the objective function is obtained by solving the adjoint equations (4.15). Each iteration of the gradient algorithm requires sequential solution of the state equation (2.7) and adjoint equation (4.15). 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 forward in time starting from the initial conditions and adjoint equations are solved marching backward in time from the final conditions at $t = T$.

Figure 2 shows the controlled velocity, temperature and concentration when the actuator of width $L = 0.1$ was placed at $x = 0.4$ on the bottom boundary. As can be seen in Figure 2 the control does a brings the
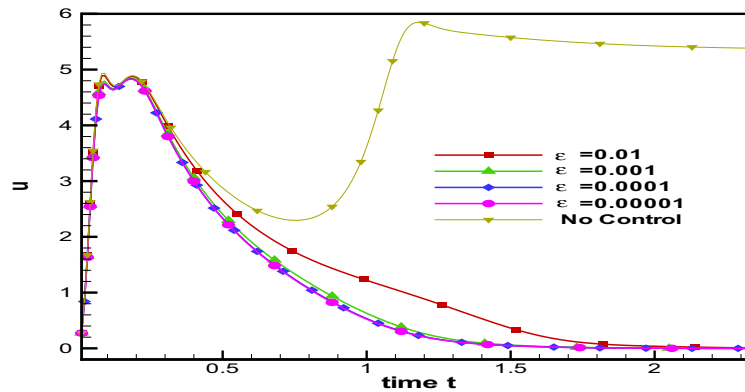
FIGURE 3. Comparison of controlled u-velocity computed for various penalty parameter ϵ. Also shown is the uncontrolled velocity. All the quantities shown are computed at $(x, y) = (3/4, 3/4)$.

convection to motion free state. As expected larger values of δ with fixed $\gamma(= 1)$ drove the convection motion close to motion free state faster albeit with large magnitude controls. The results showed that at about $T = 2$ the controlled system is close to the desired motion free state. In order to numerically verify the convergence of optimal solutions with respect to ϵ, optimal solutions were computed for a sequence of ϵ values by the optimization algorithm. Figure 3 shows the optimal u-velocity at a point inside the cavity as a function of time computed with various values of ϵ. It is clear that convergence is achieved without any oscillations as $\epsilon \to 0$.

An actuator placement study was performed to find the best place for actuation. An actuator slot of size $L = 0.1$ was placed on the bottom wall so that there is ten possible placements. Optimal solutions computed corresponding to each one of the actuator location all showed complete suppression of convection. They also showed no clear preference to actuator location although slightly faster reduction in the cost functional by the optimization algorithm was observed when the actuator was placed in the middle third part of the wall $(1/3 \le x \le 2/3)$. In particular, convergence behavior of the gradient algorithm with respect to δ and γ was also better when the actuation was on this part of the wall.

In addition to optimal placement of actuator, the size of the actuator slot can also be an issue in practical implementation of controller. We therefore studied the effect of actuator slot width on the performance of the controller. Actuator width of $L = 0.1, 0.05, 0.025, 0.0125, 0.00625$ located at $x = 0.4$ were considered on the bottom wall. As expected, as the width L decreases, it takes longer to suppress the convection, see Figure 5. The convergence of the algorithm also sensitive to the values of δ and γ with decrease in the width L. In particular, control algorithm did not converge for $\gamma = 1$ and $\delta \le 10$ when $L \le 0.0125$. This is to be expected since actuators with smaller width would require larger control to be effective in suppressing the convection.

We also investigated few other suitable choices for the cost functional in our optimal control problem. It is clear that in order to achieve our objective one can also use $\Theta_3 := \delta \|\theta - \theta_d\|^2$, $\Theta_4 := \delta \|S - S_d\|^2$, $\Theta_5 := \delta \|\theta - \theta_d\|^2 + \delta \|S - S_d\|^2$ in addition to $\Theta_1 := \delta \|\mathbf{u}\|^2 + \delta^* \|\nabla \mathbf{u}\|^2$ and $\Theta_2 : \delta \|\nabla \times \mathbf{u}\|^2$ in our control problem, where $\theta_d = S_d = 1 - x$. However, notice that only Θ_1 and Θ_2 satisfy Assumption (B), whereas Θ_3, Θ_4 and Θ_5 do not satisfy Assumption (B). Our numerical experiments showed that although all five choices perform well in suppressing the convection, the best performance were obtained with Θ_1 and Θ_2, see Figure 4. Moreover, the convergence behavior of the control algorithm is also better when Θ is chosen to satisfy Assumption (B). For example, when $\Theta = \Theta_i$, $i = 3, 4, 5$, convergence of the control algorithm was achieved only with large values of $\delta(\ge 100)$ for fixed $\gamma(= 1)$ whereas when $\Theta = \Theta_i$, $i = 1, 2$, convergence did not encounter any such limitations.
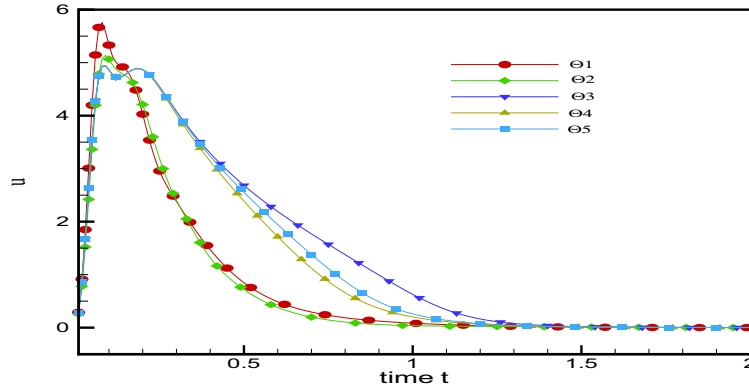
FIGURE 4. Comparison of controlled u-velocity computed with various cost functionals. All the quantities shown are computed at $(x, y) = (3/4, 3/4)$.
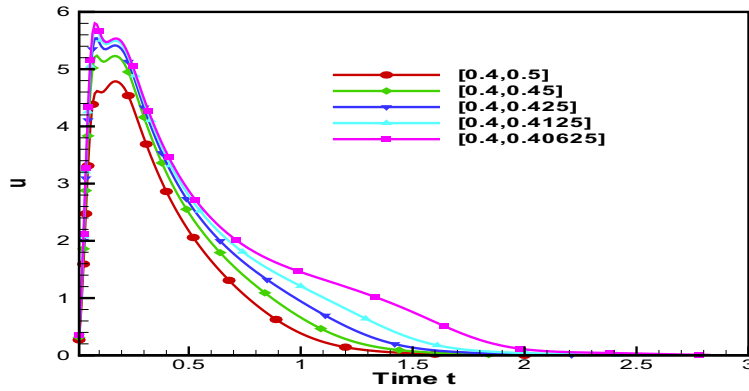

FIGURE 5. Comparison of controlled u-velocity computed with various actuator widths. All the quantities shown are computed at $(x, y) = (3/4, 3/4)$.

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 ϵ_1, $\epsilon_2 > 0$ and $\mathbf{u} \in \mathbf{V}$, $\phi, \psi \in H^1(\Omega)$, there exists a positive constant $C_{\epsilon_1, \epsilon_2}$ such that*

$$|(\mathbf{u} \cdot \nabla \phi, \psi)| \leq \epsilon_1 \|\nabla \psi\|^2 + \epsilon_2 \|\nabla \mathbf{u}\|^2 + C_{\epsilon_1, \epsilon_2} \|\mathbf{u}\|^2 \|\phi\|_{L^4(\Omega)}^4. \tag{A.1}$$

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

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

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

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\nabla \mathbf{w}\|^2 + C_\epsilon [\|\mathbf{u}\|_{L^4(\Omega)}^2 (\|\nabla \mathbf{v}\| + \|\mathbf{v}\|_{L^4(\Omega)}^2) + \|\mathbf{w}\|^2 \|\nabla \mathbf{v}\|^2]. \tag{A.3}$$

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

$$(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{u})| \leq \epsilon \|\nabla \mathbf{u}\|^2 + C_\epsilon \|\mathbf{u}\|^2 \|\nabla \mathbf{v}\|^2. \tag{A.4}$$

(v)   *For any ϵ_1, $\epsilon_2 > 0$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$, there exists a positive constant $C_{\epsilon_1, \epsilon_2}$ such that*

$$|(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{w})| \leq \epsilon_1 \|\nabla \mathbf{v}\|^2 + \epsilon_2 \|\nabla \mathbf{w}\|^2 + C_{\epsilon_1, \epsilon_2} \|\mathbf{w}\|^2 \|\mathbf{u}\|_{L^4(\Omega)}^4. \tag{A.5}$$

(vi)   *There exist constants c_1, c_2 and c_3 such that*

$$\|B_1(\mathbf{u}, \mathbf{u})\|_{4/3} \leq \|\mathbf{u}\|^{\frac{1}{2}} \|\nabla \mathbf{u}\|^{\frac{3}{2}}, \quad \|B_2(\mathbf{u}, \mathbf{u})\|_{\mathbf{V}^*} \leq \|\mathbf{u}\| \|\nabla \mathbf{u}\|, \quad and \quad \|B(\mathbf{u}, \mathbf{u})\|_{\mathbf{V}^*} \leq \|\mathbf{u}\|_1^2, \tag{A.6}$$

Proof.

(i)   First notice by Hölder's inequality

$$|(\mathbf{u} \cdot \nabla \phi, \psi)| = |(\mathbf{u} \cdot \nabla \psi, \phi)| \leq \|\mathbf{u}\|_{L^4(\Omega)} \|\phi\|_{L^4(\Omega)} \|\nabla \psi\|.$$

Now using Gagliardo−Nirenberg inequality with $q = 2$ and $r = 4$, we obtain

$$|(\mathbf{u} \cdot \nabla \phi, \psi) \leq C \|\mathbf{u}\|^{\frac{1}{2}} \|\nabla \mathbf{u}\|^{\frac{1}{2}} \|\phi\|_{L^4(\Omega)} \|\nabla \psi\|.$$

Applying Young's inequality twice, we obtain

$$|(\mathbf{u} \cdot \nabla \phi, \psi)| \leq \epsilon_1 \|\nabla \psi\|^2 + C_{\epsilon_1} \|\mathbf{u}\|^2 \|\nabla \mathbf{u}\| \|\phi\|_{L^4(\Omega)}^2$$
$$\leq \epsilon_1 \|\nabla \psi\|^2 + \epsilon_2 \|\nabla \mathbf{u}\|^2 + C_{\epsilon_1, \epsilon_2} \|\mathbf{u}\|^2 \|\phi\|_{L^4(\Omega)}^4.$$

(ii)   By definition

$$c(\mathbf{u}, \mathbf{v}, \mathbf{u}) = \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{u}) - \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v}).$$

Therefore 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.

(iii)   First by Hölder's and Gagliardo−Nirenberg's inequalities, we obtain

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

An application of Young's inequality yields

$$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq \epsilon \|\nabla \mathbf{w}\|^2 + C_\epsilon \|\mathbf{u}\|_{L^4(\Omega)}^{\frac{4}{3}} \|\nabla \mathbf{v}\|^{\frac{4}{3}} \|\mathbf{w}\|^{\frac{2}{3}} + C_\epsilon \|\mathbf{u}\|_{L^4(\Omega)}^2 \|\mathbf{v}\|_{L^4(\Omega)}^2$$
$$\leq \epsilon \|\nabla \mathbf{w}\|^2 + C_\epsilon [\|\mathbf{u}\|_{L^4(\Omega)}^2 (\|\nabla \mathbf{v}\| + \|\mathbf{v}\|_{L^4(\Omega)}^2) + \|\mathbf{w}\|^2 \|\nabla \mathbf{v}\|^2].$$

(iv)   By Hölder's and Gagliardo−Nirenberg's inequality, we obtain

$$|(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{u})| \leq \|\mathbf{u}\|_{L^4(\Omega)}^2 \|\nabla \mathbf{v}\| \leq C \|\mathbf{u}\| \|\nabla \mathbf{u}\| \|\nabla \mathbf{v}\|.$$

Now the result follows by Young's inequality.

(v)   The proof of this result is very similar to (i) and is omitted.

(vi)   Notice

$$|(B_1(\mathbf{u}, \mathbf{u}), \mathbf{v})| = |(\mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\|_{L^4(\Omega)} \|\nabla \mathbf{u}\| \|\mathbf{v}\|_{L^4(\Omega)}.$$

Therefore by Gagliardo−Nirenberg inequality

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

and the result follows from the fact that $L^{\frac{4}{3}}(\Omega)$ is the dual of $L^4(\Omega)$. The proof of the other inequalities follows by similar arguments and Sobolev imbeddings.   □

REFERENCES

[1] R.A. Adams, *Sobolev Spaces*. Academic Press, New York (1975).

[2] V.K. Andreev and I.I. Ryzhkov, Symmetry classification and exact solutions of the thermal diffusion equations. *Differ. Eqs.* **41** (2005) 538–547.

[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. *Asympotic Anal.* **71** (2011) 101–121.

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

[5] L. Bergman and M.T. Hyun, Simulation of two dimensional thermosolutal convection in liquid metals induced by horizontal temperature and species gradients. *Int. J. Heat Mass Transfer* **39** (1996) 2883.

[6] J.A. Burns, B.B. King and D. Rubio, Feedback control of thermal fluid using state estimation, Flow Control and Optimization. *Int. J. Comput. Fluid Dynamics* **11** (1998) 93–112.

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

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

[9] 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.

[10] M. Gad-el-Hak, A. Pollard and J. P. Bonnet, Flow Control, Fundamentals and Practices. *Lect. Notes Phys.* Springer, Berlin (1998).

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

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

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

[14] M.D. Gunzburger and S. Manservisi, The velocity tracking problem for Navier–Stokes flows with boundary control. *SIAM J. Control Optim.* **39** (2000) 594–634.

[15] 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.

[16] M. Hinze and K. Kunisch, Second order methods for boundary control of the instationary Navier–Stokes system. *Z. Angew. Math. Mech.* **84** (2004) 171–187.

[17] 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 and Optim.* **36** (1998) 1795–1814.

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

[19] D.T.J. Hurle and E. Jakeman, Soret driven thermo-solutal convection. *J. Fluid Mech.* **47** (1971) 667–687.

[20] K. Ito and S.S. Ravindran, Optimal control of thermally convected fluid flows. *SIAM J. Sci. Comput.* **19** (1998) 1847–1869.

[21] J.L. Lions and E. Magnes, *Problemes aux limits Non Homogeneous et Applications,* Vol. II. Dunod, Paris (1968).

[22] I. Mercader, O. Batiste, A. Alonso and E. Knoblock, Convections, anti-convections and multi-convections in binary fluid convection. *J. Fluid Mech.* **667** (2011) 586–606.

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

[24] L. Nirenberg, On elliptic partial differential equations. *Annul. Sc. Norm. Sup. Pisa* **13** (1959) 116–162.

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

[26] V.M. Shevtsova, D.E. Melnikov and J.C. Legros, Onset of convection in Soret-driven instability. *Phys. Rev. E* **73** (2006) 047302.

[27] J. Simon, Compact sets in the space $L^p(0, T; B)$ *Annali di Matematika Pura ed Applicata (IV)* **146** (1987) 65–96.

[28] J. Singer and H.H. Bau, Active control of convection. *Phys. Fluids A* **3** (1991) 2859–2865.

[29] B.L. Smorodin, Convection of a binary mixture under conditions of thermal diffusion and variable temperature gradient. *J. Appl. Mech. Tech. Phys.* **43** (2002) 217–223.

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

[31] R. Temam, Navier–Stokes Equations: Theory and Numerical Analysis. North-Holland (1977).

[32] G. Yang and N. Zabaras, The adjoint method for an inverse design problem in the directional solidification of binary alloys. *J. Comput. Phys.* **40** (1998) 432–452.