ESAIM: COCV 28 (2022) 22
https://doi.org/10.1051/cocv/2022016

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

OPTIMAL BOREL MEASURE CONTROLS FOR THE TWO-DIMENSIONAL STATIONARY BOUSSINESQ SYSTEM

Gilbert Peralta*

Abstract. We analyze an optimal control problem for the stationary two-dimensional Boussinesq system with controls taken in the space of regular Borel measures. Such measure-valued controls are known to produce sparse solutions. First-order and second-order necessary and sufficient optimality conditions are established. Following an optimize-then-discretize strategy, the corresponding finite element approximation will be solved by a semi-smooth Newton method initialized by a continuation strategy. The controls are discretized by finite linear combinations of Dirac measures concentrated at the nodes associated with the degrees of freedom for the mini-finite element.

Mathematics Subject Classification. 49J20, 49K20, 49M15.

Received June 19, 2020. Accepted February 15, 2022.

1. Introduction

Distributed controls offer advantages for optimal control problems of static and dynamic models with tracking type cost functionals. The larger the control domains, the more effective we can reach or approximate the states to a given desired target. As a result, distributed controls usually have large supports. For the past decade, there have been a large consideration in analyzing measure-valued controls due to their sparsity. Despite of the very abstract nature of measure spaces, the typical compactness arguments for reflexive Banach spaces in establishing the existence of optimal solutions can be adapted. Moreover, such control problems can be numerically solved efficiently using finite linear combinations of Dirac measures and the corresponding finite-dimensional system can be computed by a semi-smooth Newton method [5, 6].

Recently, Casas and Kunisch [8] investigated Borel measure-valued controls for the 2D stationary Naiver–Stokes equation and established first-order and second-order optimality conditions. In this paper, we will extend the study to the Boussinesq system. Due to the coupling of the Navier–Stokes and convection-diffusion equations, the appropriate functional analytic set-up must be carefully developed. In some hydrodynamical models, the variations of temperature may lead to a different fluid flow, which at the same time can affect the heat propagation due to convection. Therefore the addition of thermal controls may be a useful strategy as well. For a derivation of the Boussinesq system as an asymptotic limit of the complete Navier–Stokes system, we refer the reader to [23].

Keywords and phrases: Boussinesq system, Borel measures, sparse controls, optimality conditions, finite elements, semi-smooth Newton algorithm.

Department of Mathematics and Computer Science, University of the Philippines Baguio, Governor Pack Road, 2600 Baguio, Philippines.

* Corresponding author: grperalta@up.edu.ph

We consider the following optimal control problem

$$\min_{(\boldsymbol{\mu},\vartheta)\in \boldsymbol{M}(\omega_{\mathrm{f}})\times M(\omega_{\mathrm{t}})} J(\boldsymbol{u},\theta,\boldsymbol{\mu},\vartheta) \tag{1.1}$$

subject to the 2D Boussinesq system with homogeneous Dirichlet boundary conditions

$$\begin{cases} -\nu\Delta \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla P = \theta \boldsymbol{g} + \boldsymbol{f}_d + \chi_{\omega_{\mathrm{f}}}\boldsymbol{\mu} & \text{in } \Omega,\\ \operatorname{div} \boldsymbol{u} = 0 & \text{in } \Omega,\\ -\kappa\Delta\theta + (u\cdot\nabla)\theta = h_d + \chi_{\omega_{\mathrm{t}}}\vartheta & \text{in } \Omega,\\ \boldsymbol{u} = 0, \quad \theta = 0 & \text{on } \Gamma, \end{cases} \tag{1.2}$$

where Ω is an open and bounded C^2-domain in $\mathbb{R}^2$ with boundary Γ. In (1.2), $\boldsymbol{u} : \Omega \to \mathbb{R}^2$, $P : \Omega \to \mathbb{R}$ and $\theta : \Omega \to \mathbb{R}$ are the velocity field, pressure and temperature of the fluid. Also, $\nu > 0$ is the fluid viscosity and $\kappa > 0$ is the thermal conductivity. The control regions ω_{f} and ω_{t} are assumed to be relatively closed subsets of Ω. For controls acting only in the fluid equation one can take $\omega_{\mathrm{t}} = \emptyset$, while for controls acting only on the convection-diffusion equation one may set $\omega_{\mathrm{f}} = \emptyset$. Furthermore, $\boldsymbol{f}_d$ and h_d are external sources in the fluid and heat equations, while $\boldsymbol{g}$ is the gravitational force.

The cost functional J in (1.1) is of tracking-type given by

$$J(\boldsymbol{u},\theta,\boldsymbol{\mu},\vartheta) = \frac{1}{2}\int_\Omega |\boldsymbol{u}-\boldsymbol{u}_d|^2\,\mathrm{d}x + \frac{1}{2}\int_\Omega |\theta-\theta_d|^2\,\mathrm{d}x + \alpha\|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_{\mathrm{f}})} + \beta\|\vartheta\|_{M(\omega_{\mathrm{t}})} \tag{1.3}$$

with desired velocity and temperature $\boldsymbol{u}_d$ and θ_d, respectively. The control parameters α and β are nonnegative, with $\alpha+\beta > 0$. We denote by $M(\omega)$ the space of real and regular Borel measures on a relatively compact subset ω of Ω, and by Riesz theorem it can be identified with the dual of

$$C_0(\omega) = \{\phi \in C(\bar{\omega}) : \phi = 0 \text{ on } \partial\omega\cap\Gamma\}$$

equipped with the supremum norm $\|\phi\|_{C_0(\omega)} = \sup_{x\in\bar{\omega}}|\phi(x)|$. The associated dual norm is

$$\|\mu\|_{M(\omega)} = \sup\left\{\int_\omega \phi\,\mathrm{d}\mu : \|\phi\|_{C_0(\omega)} \le 1\right\} = |\mu|(\omega),$$

where $|\mu|$ is the total variation measure of $\mu \in M(\omega)$, see Chapter 6 of [25] for details. We equip $\boldsymbol{M}(\omega) := M(\omega)\times M(\omega)$ with the norm

$$\|(\mu_1,\mu_2)\|_{\boldsymbol{M}(\omega)} = \|\mu_1\|_{M(\omega)} + \|\mu_2\|_{M(\omega)},$$

which is the dual of $\boldsymbol{C}_0(\omega) := C_0(\omega)\times C_0(\omega)$ with the respect to the norm

$$\|(\phi_1,\phi_2)\|_{\boldsymbol{C}_0(\omega)} = \max\{\|\phi_1\|_{C_0(\omega)}, \|\phi_2\|_{C_0(\omega)}\}.$$

Due to their sparsity, regular Borel measure controls have been studied for stationary and instationary linear partial differential equations in [5, 6, 9, 10, 18, 19] and for nonlinear partial differential equations in [7, 8]. The current work follows the methodologies presented in [8], with appropriate adjustments due to the coupling with the convection-diffusion equation. The well-posedness of (1.2) is one of the crucial parts in the analysis. For this direction, we shall extend the results for the L^p-theory of (1.2) with $\boldsymbol{f}_d = \boldsymbol{0}$ and $h_d = 0$ in [17] and with $h_d = 0$ in [26], and without the measures on the right-hand sides. These can be extended to prove the existence of solutions of the Boussinesq system for source terms in Sobolev spaces with negative index and appropriate integrability. With this at hand, the extension to source terms in measure spaces is immediate thanks to the Sobolev embedding, and consequently, the existence of solutions to the above optimal control problem can be deduced from usual sequential compactness arguments. For distributed optimal controls of the Boussinesq equation in the stationary and instationary cases, see for instance [1, 4, 14, 15, 21, 22].

Let us mention some technical challenges in the analysis of the optimal control problem (1.1)–(1.3). First, the existence of very weak solutions to the Boussinesq system with rough sources can be dealt by considering a fixed velocity field on the convection terms and applying the contraction principle. Here, it is required that the data is small enough and that the powers corresponding to the integrability of the state variables lie in suitable intervals (see Lems. 2.7 and 2.8). The smallness of the data can be dropped by using perturbation, compactness and density arguments (Lem. 2.9 and Thm. 2.10). Second, as in [8] we have to verify the existence of the so-called *regular points* as a "constraint qualification" when the viscosity is large enough (Thm. 2.13). Finally, for the proof of the second-order sufficient conditions for local optimality, in particular the estimation of the convection terms appearing in the linearized state system, we need that the velocity field and the temperature to be in $\boldsymbol{W}_0^{1,p}(\Omega)$ and $W_0^{1,q}(\Omega)$, respectively, with $\frac{4}{3} < q \leq p < 2$ (see Sect. 4).

This paper will be organized as follows: In Section 2 the analysis of the state equation will be established using classical fixed point arguments and the local differentiability of the control-to-state operator will be discussed. The analysis of the optimal control problem will be the concern of Section 3, and second-order optimality conditions will be presented in Section 4. A proposed numerical scheme based on the semi-smooth Newton method and its implementation will be given in Section 5.

2. Analysis of the state equations

In this section, we prove the existence of solutions to the state equation (1.2). The usual notation for Sobolev spaces will be adapted in this paper, see [2] for instance. From the continuous embedding $W_0^{1,q}(\Omega) \subset C_0(\omega)$ for any $2 < q < \infty$, we likewise have the continuity of the embedding $M(\omega) \subset W^{-1,p}(\Omega)$ for any $1 < p < 2$ by duality. Thus, we first consider source terms that lie on the Sobolev spaces $W^{-1,p}(\Omega)$ for $1 < p < 2$. Given $1 \leq p \leq \infty$, we denote by p' the conjugate exponent of p, that is, $1/p + 1/p' = 1$. The pairing between a Banach space X and its dual X' will be denoted by $\langle f, \phi \rangle_{X' \times X}$ for $f \in X'$ and $\phi \in X$. If the space X is clear in the context, we shall simply denote this pairing by $\langle f, \phi \rangle$. For $1 < p < q < \infty$ we have $W_0^{1,q}(\Omega) \subset W_0^{1,p}(\Omega)$ and in particular $W^{-1,q}(\Omega) \subset W^{-1,p}(\Omega)$. The latter remark is also valid in the vector-valued case. Moreover, we let $\boldsymbol{L}^p(\Omega) = L^p(\Omega) \times L^p(\Omega)$, $\boldsymbol{W}^{s,p}(\Omega) = W^{s,p}(\Omega) \times W^{s,p}(\Omega)$, $\boldsymbol{W}_0^{s,p}(\Omega) = W_0^{s,p}(\Omega) \times W_0^{s,p}(\Omega)$, $\boldsymbol{H}_0^1(\Omega) = \boldsymbol{W}_0^{1,2}(\Omega)$ and $\boldsymbol{V}^p(\Omega) = \{\boldsymbol{u} \in \boldsymbol{W}_0^{1,p}(\Omega) : \operatorname{div} \boldsymbol{u} = 0 \text{ in } \Omega\}$.

2.1. Preliminaries

Let us start with the definition of very weak solutions.

Definition 2.1. *Assume that $2 < s < \infty$ and $s' < \varrho < \infty$. Suppose that $\boldsymbol{f} \in \boldsymbol{W}^{-1,s/2}(\Omega)$, $h \in W^{-1,s\varrho/(s+\varrho)}(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{L}^\infty(\Omega)$. A pair $(\boldsymbol{u}, \theta) \in \boldsymbol{L}^s(\Omega) \times L^\varrho(\Omega)$ is called a* very weak solution *of*

$$\begin{cases} -\nu \Delta \boldsymbol{u} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) + \nabla P = \theta \boldsymbol{g} + \boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div} \boldsymbol{u} = 0 & \text{in } \Omega, \\ -\kappa \Delta \theta + \operatorname{div}(\theta \boldsymbol{u}) = h & \text{in } \Omega, \\ \boldsymbol{u} = 0, \quad \theta = 0 & \text{on } \Gamma, \end{cases} \tag{2.1}$$

if the following variational equations are satisfied

$$\begin{cases} \displaystyle\int_\Omega \boldsymbol{u}\cdot(-\nu\Delta\boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi})\,\mathrm{d}x = \int_\Omega \theta\boldsymbol{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}x + \langle \boldsymbol{f},\boldsymbol{\varphi}\rangle \quad \forall\boldsymbol{\varphi}\in \boldsymbol{W}^{2,s'}(\Omega)\cap\boldsymbol{V}^{s'}(\Omega),\\ \displaystyle\int_\Omega \theta(-\kappa\Delta\psi - \boldsymbol{u}\cdot\nabla\psi)\,\mathrm{d}x = \langle h,\psi\rangle \quad \forall\psi\in W^{2,\varrho'}(\Omega)\cap W_0^{1,\varrho'}(\Omega),\\ \displaystyle\int_\Omega \boldsymbol{u}\cdot\nabla\phi\,\mathrm{d}x = 0 \quad \forall\phi\in W^{1,s'}(\Omega). \end{cases}$$

Let us verify that each integral terms in the above definition are well-defined. Indeed, from the Sobolev embedding $W_0^{1,\varrho'}(\Omega)\subset L^{2\varrho'/(2-\varrho')}(\Omega)$ and

$$\frac{1}{\varrho}+\frac{1}{s}+\frac{2-\varrho'}{2\varrho'} = \frac{1}{s}+\frac{1}{2} < 1$$

we have $\theta\boldsymbol{u}\cdot\nabla\psi\in L^1(\Omega)$ by Hölder's inequality. A similar reasoning with ϱ replaced by s yields that $\boldsymbol{u}\cdot(\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi}\in L^1(\Omega)$. Next, since $1/(s/2)' = 1-2/s = 2/s'-1 \geq (2-s')/2s'$ we have

$$\boldsymbol{W}^{2,s'}(\Omega)\cap\boldsymbol{W}_0^{1,s'}(\Omega)\subset\boldsymbol{W}_0^{1,2s'/(2-s')}(\Omega)\subset\boldsymbol{W}_0^{1,(s/2)'}(\Omega),$$

so that the duality pairing $\langle\boldsymbol{f},\boldsymbol{\varphi}\rangle$ makes sense. On the other hand, let $\eta = s\varrho/(s+\varrho)$ so that

$$\frac{1}{\eta'} = \frac{1}{\varrho'}-\frac{1}{s} > \frac{2-\varrho'}{2\varrho'}.$$

Hence, it holds that $\boldsymbol{W}^{2,\varrho'}(\Omega)\cap\boldsymbol{W}_0^{1,\varrho'}(\Omega)\subset\boldsymbol{W}_0^{1,2\varrho'/(2-\varrho')}(\Omega)\subset W_0^{1,\eta'}(\Omega)$, and as a consequence the pairing $\langle h,\psi\rangle$ is also well-defined. It is easy to see that $\theta\boldsymbol{g}\cdot\boldsymbol{\varphi}$ is integrable since $\boldsymbol{W}^{2,s'}(\Omega)\subset\boldsymbol{W}^{1,2s'/(2-s')}(\Omega)\subset\boldsymbol{W}^{1,2}(\Omega)\subset\boldsymbol{L}^p(\Omega)$ for any $1\leq p<\infty$. Finally, the remaining terms can be checked to be well-defined as well.

To recover the pressure from the first equation in the very weak formulation (2.1), we invoke de Rham's theorem [12] to conclude the existence of $P\in W^{-1,s}(\Omega)$ such that

$$-\nu\Delta\boldsymbol{u}+\operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u})+\nabla P = \theta\boldsymbol{g}+\boldsymbol{f} \text{ in } \boldsymbol{W}^{-2,s}(\Omega) \tag{2.2}$$

in the sense of distributions. The third equation in (2.1) implies the incompressibility property

$$\operatorname{div}\boldsymbol{u} = 0 \text{ in } W^{1,s'}(\Omega)'. \tag{2.3}$$

Finally, from the second equation in (2.1), we have

$$-\kappa\Delta\theta+\operatorname{div}(\theta\boldsymbol{u}) = h \text{ in } W^{-2,\varrho}(\Omega). \tag{2.4}$$

With regard to the boundary conditions, we recall standard results for the very weak solutions to the Stokes and Poisson equations in [16] and the regularity of the corresponding generalized trace on the boundary. In particular, this implies that the boundary conditions in (2.1) are satisfied in the sense of generalized traces stipulated below.

Theorem 2.2 ([16], Thm. 5)**.** *Let $1<q<\infty$ and*

$$\boldsymbol{H} = \{\boldsymbol{u}\in\boldsymbol{L}^q(\Omega) : \operatorname{div}\boldsymbol{u}\in W^{1,q'}(\Omega)',\ -\Delta\boldsymbol{u}+\nabla P\in(\boldsymbol{W}^{2,q'}(\Omega)\cap\boldsymbol{W}_0^{1,q'}(\Omega))' \text{ for some } P\in W^{-1,q}(\Omega)\}.$$

Then there exists a unique trace operator $\boldsymbol{\gamma} : \boldsymbol{H} \to \boldsymbol{W}^{-1/q,q}(\Gamma)$ *such that*

$$\langle \boldsymbol{\gamma u}, \partial_N \boldsymbol{\varphi} \rangle = \int_\Omega \boldsymbol{u} \cdot \Delta \boldsymbol{\varphi} \,\mathrm{d}x + \langle -\Delta \boldsymbol{u} + \nabla P, \boldsymbol{\varphi} \rangle$$

for every $\boldsymbol{\varphi} \in \boldsymbol{W}^{2,q'}(\Omega) \cap \boldsymbol{W}_0^{1,q'}(\Omega)$ *and*

$$\langle \boldsymbol{\gamma u}, \phi N \rangle = \int_\Omega \boldsymbol{u} \cdot \nabla \phi \,\mathrm{d}x + \langle \operatorname{div} \boldsymbol{u}, \phi \rangle$$

for every $\phi \in W^{1,q'}(\Omega)$*. Furthermore, there is a constant* $c = c_{q,\Omega} > 0$ *independent on* $\boldsymbol{u}$ *such that*

$$\|\boldsymbol{\gamma u}\|_{\boldsymbol{W}^{-1/q,q}(\Gamma)} \leq c\{\|\boldsymbol{u}\|_{\boldsymbol{L}^q(\Omega)} + \|\operatorname{div} \boldsymbol{u}\|_{W^{1,q'}(\Omega)'} + \|-\Delta \boldsymbol{u} + \nabla P\|_{(\boldsymbol{W}^{2,q'}(\Omega) \cap \boldsymbol{W}_0^{1,q'}(\Omega))'}\}.$$

If $\boldsymbol{u} \in \boldsymbol{W}^{1,q}(\Omega)$ *and* $P \in L^q(\Omega)$*, then* $\gamma \boldsymbol{u}$ *coincides with the usual trace of* $\boldsymbol{u}$ *on* Γ*.*

Theorem 2.3 ([16], Thm. 6)**.** *Let* $1 < q < \infty$ *and*

$$H = \{\theta \in L^q(\Omega) : \Delta\theta \in (W^{2,q'}(\Omega) \cap W_0^{1,q'}(\Omega))'\}.$$

Then there exists a unique trace operator $\gamma : H \to W^{-1/q,q}(\Gamma)$ *such that*

$$\langle \gamma\theta, \partial_N \psi \rangle = \int_\Omega \theta \Delta \psi \,\mathrm{d}x - \langle \Delta\theta, \psi \rangle$$

for every $\psi \in W^{2,q'}(\Omega) \cap W_0^{1,q'}(\Omega)$ *and there exists a constant* $c = c_{q,\Omega} > 0$ *such that*

$$\|\gamma\theta\|_{W^{-1/q,q}(\Gamma)} \leq c\{\|\theta\|_{L^q(\Omega)} + \|\Delta\theta\|_{(W^{2,q'}(\Omega) \cap W_0^{1,q'}(\Omega))'}\}.$$

If $\theta \in W^{1,q}(\Omega)$ *then* $\gamma\theta$ *coincides with the usual trace of* θ *on* Γ*.*

As usual, we drop the trace operator in the notation. For instance, $\theta = 0$ on Γ precisely means that $\gamma\theta = 0$ in $W^{-1/q,q}(\Omega)$ in the sense of the previous theorem. The following theorem justifies the notion of very weak solutions presented above.

Theorem 2.4. *If* $(\boldsymbol{u}, \theta) \in \boldsymbol{L}^s(\Omega) \times L^\varrho(\Omega)$*, with* s *and* ϱ *as stated in Definition 2.1, is a very weak solution of* (2.1) *then* (2.2)–(2.4) *are satisfied for some* $P \in W^{-1,s}(\Omega)$*, and moreover,* $\boldsymbol{u} = \boldsymbol{0}$ *in* $\boldsymbol{W}^{-1/s,s}(\Gamma)$ *and* $\theta = 0$ *in* $W^{-1/\varrho,\varrho}(\Gamma)$*.*

Proof. This is a direct consequence of the above discussions along with Theorem 2.2 and Theorem 2.3. □

We close this subsection by stating the existence and uniqueness of very weak solutions to the Stokes and Poisson equations, whose proofs can be found in [16]. These will be useful when we decouple the fluid and convection-diffusion equations in the fixed point arguments. We would like to point out that the definitions of very weak solutions to (2.5) and (2.6) below are analogous to the formulation given in Definition 2.1.

Theorem 2.5 ([16], Thms. 4 and 7)**.** *Let* $q, r, s \in (1, \infty)$ *with* $s \geq 2q/(2+q)$*. Consider the Stokes problem*

$$\begin{cases} -\nu \Delta \boldsymbol{y} + \nabla \pi = \boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div} \boldsymbol{y} = 0 & \text{in } \Omega, \\ \boldsymbol{y} = 0 & \text{on } \Gamma. \end{cases} \tag{2.5}$$

If $\boldsymbol{f} \in \boldsymbol{L}^r(\Omega) + \boldsymbol{W}^{-1,s}(\Omega)$*, then* (2.5) *has a unique very weak solution* $\boldsymbol{y} \in \boldsymbol{L}^q(\Omega)$ *and there exists a constant* $c_S = c_S(q, r, s, \Omega) > 0$ *such that*

$$\|\boldsymbol{y}\|_{\boldsymbol{L}^q(\Omega)} \leq \frac{c_S}{\nu}\|\boldsymbol{f}\|_{\boldsymbol{L}^r(\Omega)+\boldsymbol{W}^{-1,s}(\Omega)}.$$

In addition, if $\boldsymbol{f} \in \boldsymbol{W}^{-1,q}(\Omega)$ *then the very weak solution satisfies* $\boldsymbol{y} \in \boldsymbol{V}^q(\Omega)$ *and there exists* $c_S = c_S(q, \Omega) > 0$ *such that*

$$\|\boldsymbol{y}\|_{\boldsymbol{V}^q(\Omega)} \leq \frac{c_S}{\nu}\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,q}(\Omega)}.$$

Theorem 2.6 ([16], Thm. 9)**.** *Let* $r, s \in (1, \infty)$ *be such that* $s \geq 2r/(2 + r)$ *and consider the Poisson equation*

$$\begin{cases} -\kappa\Delta\theta = h \ \text{ in } \Omega, \\ \qquad \theta = 0 \ \text{ on } \Gamma. \end{cases} \tag{2.6}$$

If $h \in W^{-1,s}(\Omega)$ *then* (2.6) *has a unique very weak solution* $\theta \in L^r(\Omega)$ *and there is a constant* $c_P = c_P(r, s, \Omega) > 0$ *such that*

$$\|\theta\|_{L^r(\Omega)} \leq \frac{c_P}{\kappa}\|h\|_{W^{-1,s}(\Omega)}.$$

In addition, if $h \in W^{-1,r}(\Omega)$ *then* $\theta \in W^{1,r}_0(\Omega)$ *and there exists* $c_P = c_P(r, \Omega) > 0$ *such that*

$$\|\theta\|_{W^{1,r}_0(\Omega)} \leq \frac{c_P}{\kappa}\|h\|_{W^{-1,r}(\Omega)}.$$

2.2. Well-posedness of the state equation

We begin with a lemma concerning the solutions of the convection-diffusion equation with a given velocity field. For each $\delta > 0$, $\boldsymbol{B}_{p,\delta}$ denote the closed ball in $\boldsymbol{L}^p(\Omega)$ centered at the origin having a radius δ.

Lemma 2.7. *Let* $2 < s < \infty$ *and* $s' < \varrho < \infty$*. There exists* $\delta_0 = \delta_0(s, \varrho, \Omega) > 0$ *such that for every* $h \in W^{-1,\varrho s/(\varrho+s)}(\Omega)$ *and* $\boldsymbol{u} \in \boldsymbol{B}_{s,\kappa\delta_0}$ *the convection-diffusion equation*

$$\begin{cases} -\kappa\Delta\theta_{\boldsymbol{u}} + \operatorname{div}(\theta_{\boldsymbol{u}}\boldsymbol{u}) = h \ \text{ in } \Omega, \\ \qquad\qquad\qquad\quad \theta_{\boldsymbol{u}} = 0 \ \text{ on } \Gamma, \end{cases} \tag{2.7}$$

has a unique very weak solution $\theta_{\boldsymbol{u}} \in L^\varrho(\Omega)$*. Furthermore, there exists a constant* $c_0 = c_0(s, \varrho, \Omega) > 0$ *such that for every* $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{B}_{s,\kappa\delta_0}$ *we have*

$$\|\theta_{\boldsymbol{u}}\|_{L^\varrho(\Omega)} \leq \frac{c_0}{\kappa}\|h\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)}, \tag{2.8}$$

$$\|\theta_{\boldsymbol{u}} - \theta_{\boldsymbol{v}}\|_{L^\varrho(\Omega)} \leq \frac{c_0^2}{\kappa^2}\|h\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)}\|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)}. \tag{2.9}$$

Proof. We adapt the proof based on the contraction principle in Lemma 3 of [17]. Note that given $\theta \in L^\varrho(\Omega)$ and $\boldsymbol{u} \in \boldsymbol{L}^s(\Omega)$ we have $\operatorname{div}(\theta\boldsymbol{u}) \in W^{-1,s\varrho/(s+\varrho)}(\Omega)$. From Theorem 2.6 and the fact that $\varrho s/(\varrho + s) > 2\varrho/(\varrho + 2)$, the Poisson problem

$$\begin{cases} -\kappa\Delta\widehat{\theta} = h - \operatorname{div}(\theta \boldsymbol{u}) & \text{in } \Omega, \\ \widehat{\theta} = 0 & \text{on } \Gamma, \end{cases}$$

admits a unique very weak solution $\widehat{\theta} \in L^{\varrho}(\Omega)$ and by Hölder's inequality

$$\begin{aligned} \|\widehat{\theta}\|_{L^{\varrho}(\Omega)} &\leq \frac{c_P}{\kappa}\|h - \operatorname{div}(\theta\boldsymbol{u})\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)} \\ &\leq \frac{c_P}{\kappa}\|h\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)} + \frac{c_P}{\kappa}\|\boldsymbol{u}\|_{\boldsymbol{L}^s(\Omega)}\|\theta\|_{L^{\varrho}(\Omega)}. \end{aligned}$$

Let us define the map $S : L^{\varrho}(\Omega) \to L^{\varrho}(\Omega)$ by $S(\theta) = \widehat{\theta}$. Then for $\boldsymbol{u} \in \boldsymbol{B}_{s,\kappa\delta_0}$ we have the Lipschitz estimate

$$\|S(\theta_1) - S(\theta_2)\|_{L^{\varrho}(\Omega)} \leq \frac{c_P}{\kappa}\|\boldsymbol{u}\|_{\boldsymbol{L}^s(\Omega)}\|\theta_1 - \theta_2\|_{L^{\varrho}(\Omega)} \leq c_P\delta_0\|\theta_1 - \theta_2\|_{L^{\varrho}(\Omega)}$$

for any $\theta_1, \theta_2 \in L^{\varrho}(\Omega)$. Choosing $\delta_0 > 0$ such that $c_P\delta_0 < 1/2$, we can see that S is a contraction. Therefore the boundary value problem (2.7) admits a unique very weak solution $\theta_{\boldsymbol{u}} \in L^{\varrho}(\Omega)$ and moreover the estimate (2.8) with $c_0 = 2c_P$ follows from the choice of δ_0. With regards to the stability with respect to the velocity field we have

$$\begin{aligned} \|\theta_{\boldsymbol{u}} - \theta_{\boldsymbol{v}}\|_{L^{\varrho}(\Omega)} &\leq \frac{c_P}{\kappa}\|\theta_{\boldsymbol{u}}\boldsymbol{u} - \theta_{\boldsymbol{v}}\boldsymbol{v}\|_{\boldsymbol{W}^{-1,\varrho s/(\varrho+s)}(\Omega)} \\ &\leq \frac{c_P}{\kappa}\|\boldsymbol{u}\|_{\boldsymbol{L}^s(\Omega)}\|\theta_{\boldsymbol{u}} - \theta_{\boldsymbol{v}}\|_{L^{\varrho}(\Omega)} + \frac{c_P}{\kappa}\|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)}\|\theta_{\boldsymbol{v}}\|_{L^{\varrho}(\Omega)} \\ &\leq \frac{1}{2}\|\theta_{\boldsymbol{u}} - \theta_{\boldsymbol{v}}\|_{L^{\varrho}(\Omega)} + \frac{2c_P^2}{\kappa^2}\|h\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)}\|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)} \end{aligned}$$

and after rearrangement this yields the second estimate (2.9). □

Next, we prove the existence of very weak solutions to (2.1) under a smallness condition on the data.

Lemma 2.8. *Let $2 < s < \infty$ and $s' < \varrho < \infty$. Assume that $\boldsymbol{f} \in \boldsymbol{W}^{-1,s/2}(\Omega)$, $h \in W^{-1,s\varrho/(s+\varrho)}(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{L}^{\infty}(\Omega)$. There exists $\delta_1 = \delta_1(s, \varrho, \Omega) > 0$ such that if*

$$\frac{1}{\nu\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^{\infty}(\Omega)}\|h\|_{W^{-1,s\varrho/(\varrho+s)}(\Omega)} + \frac{1}{\nu}\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,s/2}(\Omega)} \leq \delta_1 \tag{2.10}$$

then (2.1) has a very weak solution $(\boldsymbol{u}, \theta) \in \boldsymbol{L}^s(\Omega) \times L^{\varrho}(\Omega)$. Furthermore, there is a constant $c_1 = c_1(s, \varrho, \Omega) > 0$ independent on $\boldsymbol{u}$ and θ such that

$$\|\theta\|_{L^{\varrho}(\Omega)} \leq \frac{c_1}{\kappa}\|h\|_{W^{-1,\varrho s/(\varrho+s)}(\Omega)} \tag{2.11}$$

$$\|\boldsymbol{u}\|_{\boldsymbol{L}^s(\Omega)} \leq \frac{c_1}{\nu}\left(\frac{1}{\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^{\infty}(\Omega)}\|h\|_{W^{-1,s\varrho/(s+\varrho)}(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,s/2}(\Omega)}\right). \tag{2.12}$$

Proof. Given $\boldsymbol{u} \in \boldsymbol{B}_{s,\kappa\delta_0}$, let $\theta_{\boldsymbol{u}} \in L^{\varrho}(\Omega)$ be the solution of (2.7). Let us consider the Stokes equation

$$\begin{cases} -\nu\Delta\boldsymbol{y}_{\boldsymbol{u}} + \nabla\pi = \theta_{\boldsymbol{u}}\boldsymbol{g} + \boldsymbol{f} - \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{y}_{\boldsymbol{u}} = 0 & \text{in } \Omega, \\ \boldsymbol{y}_{\boldsymbol{u}} = 0 & \text{on } \Gamma. \end{cases} \tag{2.13}$$

Since $s > 2$ we have $s/2 > 2s/(2+s)$. Also, because $\theta_{\boldsymbol{u}}\boldsymbol{g} \in \boldsymbol{L}^{\varrho}(\Omega)$ and $\boldsymbol{f} - \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) \in \boldsymbol{W}^{-1,s/2}(\Omega)$, it follows from Theorem 2.5 that the system (2.13) has a unique very weak solution $\boldsymbol{y}_{\boldsymbol{u}} \in \boldsymbol{L}^s(\Omega)$, and moreover, by Lemma 2.7 and Hölder's inequality we have

$$\begin{aligned}\|\boldsymbol{y}_{\boldsymbol{u}}\|_{\boldsymbol{L}^s(\Omega)} &\le \frac{c_S}{\nu}\left(\|\theta_{\boldsymbol{u}}\boldsymbol{g}\|_{\boldsymbol{L}^{\varrho}(\Omega)} + \|\boldsymbol{u}\otimes\boldsymbol{u}\|_{\boldsymbol{L}^{s/2}(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,s/2}(\Omega)}\right)\\ &\le \frac{c_S}{\nu}\left(\frac{c_0}{\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|h\|_{W^{-1,s\varrho/(s+\varrho)}(\Omega)} + \|\boldsymbol{u}\|^2_{\boldsymbol{L}^s(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,s/2}(\Omega)}\right).\end{aligned} \tag{2.14}$$

Therefore the map $\boldsymbol{S}_\delta : \boldsymbol{B}_{s,\kappa\delta} \to \boldsymbol{L}^s(\Omega)$ given by $\boldsymbol{S}_\delta(\boldsymbol{u}) = \boldsymbol{y}_{\boldsymbol{u}}$ is well-defined for every $0 < \delta < \delta_0$, where $\delta_0 > 0$ is the constant in Lemma 2.7.

Let us show that for some $0 < \delta < \delta_0$, $\boldsymbol{S}_\delta$ maps the ball $\boldsymbol{B}_{s,\kappa\delta}$ into itself. Indeed, one may take $\delta < \min(1, \delta_0, \frac{\nu}{2c_S\kappa\max\{c_0,1\}})$ and then choose $0 < \delta_1 < \frac{1}{\nu}\kappa^2\delta^2$. If $\boldsymbol{u} \in \boldsymbol{B}_{s,\kappa\delta}$ then

$$\|\boldsymbol{S}_\delta(\boldsymbol{u})\|_{\boldsymbol{L}^s(\Omega)} \le c_S\max\{c_0,1\}\left(\frac{1}{\nu}\kappa^2\delta^2 + \delta_1\right) < \frac{2c_S}{\nu}\max\{c_0,1\}\kappa^2\delta^2 < \kappa\delta$$

according to (2.10) and (2.14). Hence, it holds that $\boldsymbol{S}_\delta(\boldsymbol{B}_{s,\kappa\delta}) \subset \boldsymbol{B}_{s,\kappa\delta}$. Moreover, if we further reduce δ and δ_1 so that

$$\frac{c_S}{\nu}\left(2\kappa\delta + \frac{\delta_1\nu c_0^2}{\kappa}\right) < 1,$$

then $\boldsymbol{S}_\delta$ is a contraction map. Indeed, for $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{B}_{s,\kappa\delta}$ we have

$$\begin{aligned}&\|\boldsymbol{S}_\delta(\boldsymbol{u}) - \boldsymbol{S}_\delta(\boldsymbol{v})\|_{\boldsymbol{L}^s(\Omega)}\\ &\quad\le \frac{c_S}{\nu}\left(\|(\boldsymbol{u}-\boldsymbol{v})\otimes\boldsymbol{u}\|_{\boldsymbol{L}^{s/2}(\Omega)} + \|\boldsymbol{v}\otimes(\boldsymbol{u}-\boldsymbol{v})\|_{\boldsymbol{L}^{s/2}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\theta_{\boldsymbol{u}} - \theta_{\boldsymbol{v}}\|_{L^{\varrho}(\Omega)}\right)\\ &\quad\le \frac{c_S}{\nu}\left(\|\boldsymbol{u}\|_{\boldsymbol{L}^s(\Omega)} + \|\boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)} + \frac{c_0^2}{\kappa^2}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|h\|_{W^{-1,s\varrho/(s+\varrho)}(\Omega)}\right)\|\boldsymbol{u}-\boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)}\\ &\quad\le \frac{c_S}{\nu}\left(2\kappa\delta + \frac{\delta_1\nu c_0^2}{\kappa}\right)\|\boldsymbol{u}-\boldsymbol{v}\|_{\boldsymbol{L}^s(\Omega)}.\end{aligned}$$

Hence, by the Banach contraction principle, $\boldsymbol{S}_\delta$ has a unique fixed point, which corresponds to a very weak solution of the system (2.1). The estimates (2.11) and (2.12) can be deduced from (2.8) and (2.14), reducing $\delta > 0$ if necessary. □

The second step is to establish the existence of weak solutions for a perturbed Boussinesq system.

Lemma 2.9. *Suppose that* $2 < \gamma < \infty$, $2 < \beta < \infty$, $\boldsymbol{g} \in \boldsymbol{L}^\infty(\Omega)$, $\boldsymbol{v} \in \boldsymbol{L}^\gamma(\Omega)$, $\eta \in L^\beta(\Omega)$, $\boldsymbol{f} \in \boldsymbol{H}^{-1}(\Omega)$ *and* $h \in H^{-1}(\Omega)$ *with* $\operatorname{div}\boldsymbol{v} = 0$. *There exists* $\delta_2 = \delta_2(\gamma,\beta,\Omega) > 0$ *such that if*

$$\frac{1}{\nu\kappa}\|\eta\|_{L^\beta(\Omega)} + \frac{1}{\nu}\|\boldsymbol{v}\|_{\boldsymbol{L}^\gamma(\Omega)} \le \delta_2 \tag{2.15}$$

then the boundary value problem

$$\begin{cases} -\nu\Delta\boldsymbol{u} + \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u} + \boldsymbol{u}\otimes\boldsymbol{v} + \boldsymbol{v}\otimes\boldsymbol{u}) + \nabla\pi = \theta\boldsymbol{g} + \boldsymbol{f} & \text{in } \Omega,\\ \operatorname{div}\boldsymbol{u} = 0 & \text{in } \Omega,\\ -\kappa\Delta\theta + \operatorname{div}(\theta\boldsymbol{u} + \theta\boldsymbol{v} + \eta\boldsymbol{u}) = h & \text{in } \Omega,\\ \boldsymbol{u} = 0, \quad \theta = 0 & \text{on } \Gamma,\end{cases} \tag{2.16}$$

has at least one weak solution $(\boldsymbol{u},\theta) \in \boldsymbol{V}^2(\Omega) \times H_0^1(\Omega)$.

Proof. The proof is again based on a fixed point argument. Consider the map $S : \boldsymbol{V}^2(\Omega) \times H_0^1(\Omega) \to \boldsymbol{V}^2(\Omega) \times H_0^1(\Omega)$ defined by $S(\bar{\boldsymbol{u}},\bar\theta) = (\boldsymbol{u},\theta)$, where $\boldsymbol{u} \in \boldsymbol{V}^2(\Omega)$ is the weak solution of the linearized Stokes equation

$$\begin{cases} -\nu\Delta \boldsymbol{u} + \nabla\pi = \theta \boldsymbol{g} + \boldsymbol{f} - \operatorname{div}(\bar{\boldsymbol{u}}\otimes\bar{\boldsymbol{u}} + \bar{\boldsymbol{u}}\otimes\boldsymbol{v} + \boldsymbol{v}\otimes\bar{\boldsymbol{u}}) & \text{in } \Omega,\\ \qquad\operatorname{div}\boldsymbol{u} = 0 & \text{in } \Omega,\\ \qquad\quad \boldsymbol{u} = 0 & \text{on } \Gamma, \end{cases} \tag{2.17}$$

while $\theta \in H_0^1(\Omega)$ is the weak solution of the Poisson equation

$$\begin{cases} -\kappa\Delta\theta = h - \operatorname{div}(\bar\theta\bar{\boldsymbol{u}} + \bar\theta\boldsymbol{v} + \eta\bar{\boldsymbol{u}}) & \text{in } \Omega,\\ \qquad\theta = 0 & \text{on } \Gamma. \end{cases} \tag{2.18}$$

Observe that the right-hand sides of the above equations belong to $\boldsymbol{H}^{-1}(\Omega)$ and $H^{-1}(\Omega)$, respectively, hence existence and uniqueness of weak solutions are guaranteed according to classical elliptic theory.

Let us show that S is compact. Suppose that $(\bar{\boldsymbol{u}}_n,\bar\theta_n)$ is a bounded sequence in $\boldsymbol{V}^2(\Omega)\times H_0^1(\Omega)$ so that up to a subsequence, which we do not relabel for simplicity, we have $\bar{\boldsymbol{u}}_n \rightharpoonup \bar{\boldsymbol{u}}$ in $\boldsymbol{V}^2(\Omega)$ and $\bar\theta_n \rightharpoonup \bar\theta$ in $H_0^1(\Omega)$. By the compactness of the embedding $\boldsymbol{V}^2(\Omega)\times H_0^1(\Omega) \subset [\boldsymbol{L}^{\beta^*}(\Omega)\cap\boldsymbol{L}^{\gamma^*}(\Omega)]\times L^{\gamma^*}(\Omega)$, where $1/\gamma + 1/\gamma^* = 1/2$ and $1/\beta + 1/\beta^* = 1/2$, we can further extract a subsequence in such a way that $\bar{\boldsymbol{u}}_n \to \bar{\boldsymbol{u}}$ in $\boldsymbol{L}^{\beta^*}(\Omega)\cap\boldsymbol{L}^{\gamma^*}(\Omega)$ and $\bar\theta_n \to \bar\theta$ in $L^{\gamma^*}(\Omega)$. Let $S(\bar{\boldsymbol{u}}_n,\bar\theta_n) = (\boldsymbol{u}_n,\theta_n)$ and $S(\bar{\boldsymbol{u}},\bar\theta) = (\boldsymbol{u},\theta)$. According to the standard a priori error estimate for the Poisson equation we have

$$\begin{aligned} \|\theta_n - \theta\|_{H_0^1(\Omega)} &\le \frac{c}{\kappa}(\|\bar\theta_n\bar{\boldsymbol{u}}_n + \bar\theta_n\boldsymbol{v} + \eta\bar{\boldsymbol{u}}_n - \bar\theta\bar{\boldsymbol{u}} - \bar\theta\boldsymbol{v} - \eta\bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)})\\ &= \frac{c}{\kappa}(\|\bar\theta_n(\bar{\boldsymbol{u}}_n - \bar{\boldsymbol{u}}) + (\bar\theta_n - \bar\theta)\boldsymbol{v} + \eta(\bar{\boldsymbol{u}}_n - \bar{\boldsymbol{u}}) + (\bar\theta_n - \bar\theta)\bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)})\\ &\le \frac{c}{\kappa}\{(\|\bar\theta_n\|_{H_0^1(\Omega)} + \|\eta\|_{L^\beta(\Omega)})\|\bar{\boldsymbol{u}}_n - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^{\beta^*}(\Omega)} + \|\bar\theta_n - \bar\theta\|_{L^{\gamma^*}(\Omega)}(\|\bar{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)} + \|\boldsymbol{v}\|_{\boldsymbol{L}^\gamma(\Omega)})\}, \end{aligned}$$

where we used the continuity of the embedding $\boldsymbol{V}^2(\Omega)\subset\boldsymbol{L}^\gamma(\Omega)$ and $H_0^1(\Omega)\subset L^\beta(\Omega)$ in the last step. Therefore, we have the strong convergence $\theta_n \to \theta$ in $H_0^1(\Omega)$. Similarly, since $\bar{\boldsymbol{u}}_n \to \bar{\boldsymbol{u}}$ in $\boldsymbol{L}^{\beta^*}(\Omega)\cap\boldsymbol{L}^{\gamma^*}(\Omega)$ and

$$\begin{aligned} \|\boldsymbol{u}_n - \boldsymbol{u}\|_{\boldsymbol{V}^2(\Omega)} &\le \frac{c}{\nu}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\theta_n - \theta\|_{L^2(\Omega)}\\ &\quad + \frac{c}{\nu}\Big[(\|\bar{\boldsymbol{u}}_n\|_{\boldsymbol{V}^2(\Omega)} + \|\bar{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)})\|\bar{\boldsymbol{u}}_n - \bar{\boldsymbol{u}}\|_{L^{\beta^*}(\Omega)} + 2\|\boldsymbol{v}\|_{\boldsymbol{L}^\gamma(\Omega)}\|\bar{\boldsymbol{u}}_n - \bar{\boldsymbol{u}}\|_{L^{\gamma^*}(\Omega)}\Big] \end{aligned}$$

we have $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\boldsymbol{V}^2(\Omega)$. Therefore S is a compact operator

The next step is to show that set of all possible fixed points of αS for $0 \le \alpha \le 1$ is uniformly bounded. Suppose that $(\boldsymbol{u}_\alpha,\theta_\alpha) = \alpha S(\boldsymbol{u}_\alpha,\theta_\alpha)$ and $(\tilde{\boldsymbol{u}},\tilde\theta) = S(\boldsymbol{u}_\alpha,\theta_\alpha)$. Therefore, $(\tilde{\boldsymbol{u}},\tilde\theta)\in\boldsymbol{V}^2(\Omega)\times H_0^1(\Omega)$ satisfies the following system

$$\begin{cases} -\nu\Delta\tilde{\boldsymbol{u}} + \nabla\tilde\pi = \tilde\theta\boldsymbol{g} + \boldsymbol{f} - \alpha\operatorname{div}(\alpha(\tilde{\boldsymbol{u}}\otimes\tilde{\boldsymbol{u}}) + \tilde{\boldsymbol{u}}\otimes\boldsymbol{v} + \boldsymbol{v}\otimes\tilde{\boldsymbol{u}}) & \text{in } \Omega,\\ \qquad\operatorname{div}\tilde{\boldsymbol{u}} = 0 & \text{in } \Omega,\\ \quad -\kappa\Delta\tilde\theta = h - \alpha\operatorname{div}(\alpha\tilde\theta\tilde{\boldsymbol{u}} + \tilde\theta\boldsymbol{v} + \eta\tilde{\boldsymbol{u}}) & \text{in } \Omega,\\ \qquad\quad\tilde{\boldsymbol{u}} = 0, \qquad \tilde\theta = 0 & \text{on } \Gamma, \end{cases} \tag{2.19}$$

since $\alpha(\tilde{\boldsymbol{u}},\tilde{\theta}) = (\boldsymbol{u}_\alpha,\theta_\alpha)$. Using the test function $\tilde{\theta}$ on the third equation in (2.19) along with $\operatorname{div}\boldsymbol{v} = \operatorname{div}\tilde{\boldsymbol{u}} = 0$ we have

$$\|\tilde{\theta}\|_{H_0^1(\Omega)} \le \frac{c}{\kappa}(\|h\|_{W^{-1,2}(\Omega)} + \alpha\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)}\|\eta\|_{L^\beta(\Omega)}). \tag{2.20}$$

On the other hand, applying the test function $\tilde{\boldsymbol{u}}$ to the first equation in (2.19) and then using (2.20) yield

$$\begin{aligned}
\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)} &\le \frac{c}{\nu}(\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,2}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\tilde{\theta}\|_{H_0^1(\Omega)} + \alpha\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{L}^\gamma(\Omega)})\\
&\le \frac{c}{\nu}\Big[\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,2}(\Omega)} + \frac{c}{\kappa}\|\boldsymbol{g}\|_{L^\infty(\Omega)}\left(\|h\|_{W^{-1,2}(\Omega)} + \alpha\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)}\|\eta\|_{L^\beta(\Omega)}\right)\\
&\qquad + \alpha\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}^2(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{L}^\gamma(\Omega)}\Big].
\end{aligned} \tag{2.21}$$

Therefore, choosing $\delta_2 > 0$ small enough, and recalling that $(\boldsymbol{u}_\alpha,\theta_\alpha) = \alpha(\tilde{\boldsymbol{u}},\tilde{\theta})$ and $0 \le \alpha \le 1$, we obtain from (2.20) and (2.21) that

$$\|\boldsymbol{u}_\alpha\|_{\boldsymbol{V}^2(\Omega)} + \|\theta_\alpha\|_{H_0^1(\Omega)} \le \frac{c}{\nu}\left(\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,2}(\Omega)} + \frac{1}{\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|h\|_{W^{-1,2}(\Omega)}\right).$$

By the Leray–Schauder Theorem, S has a fixed point which corresponds to a weak solution of the perturbed Boussinesq equation (2.16). □

Now we are in a position to prove the existence of very weak solutions without the smallness condition on the data. If on the other hand this smallness condition is reinforced, we also have the uniqueness and stability of the very weak solution. We only restrict to the case where $2 < \varrho < \infty$, for which more regularity of the very weak solution is available.

Theorem 2.10. *Let $2 < s < \infty$ and $2 < \varrho < \infty$. Assume that $\boldsymbol{f} \in \boldsymbol{W}^{-1,s/2}(\Omega)$, $h \in W^{-1,s\varrho/(s+\varrho)}(\Omega)$ and $\boldsymbol{g} \in L^\infty(\Omega)$. Then (2.1) has a very weak solution $(\boldsymbol{u},\theta) \in \boldsymbol{L}^s(\Omega) \times L^\varrho(\Omega)$. Moreover, there exists P such that the triple*

$$(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^\tau(\Omega) \times (L^\tau(\Omega)/\mathbb{R}) \times W_0^{1,s\varrho/(s+\varrho)}(\Omega),$$

where $\tau = \min(\varrho, s/2)$, is a solution of the following variational system

$$\begin{cases}
\displaystyle\int_\Omega \{\nu\nabla\boldsymbol{u}\cdot\nabla\boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi}\cdot\boldsymbol{u}\}\,\mathrm{d}x - \int_\Omega P\operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x = \int_\Omega \theta\boldsymbol{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}x + \langle\boldsymbol{f},\boldsymbol{\varphi}\rangle \quad \forall\boldsymbol{\varphi} \in \boldsymbol{W}_0^{1,\tau'}(\Omega),\\
\displaystyle\int_\Omega \{\kappa\nabla\theta\cdot\nabla\psi - \theta\boldsymbol{u}\cdot\nabla\psi\}\,\mathrm{d}x = \langle h,\psi\rangle \quad \forall\psi \in W_0^{1,(\varrho s/(s+\varrho))'}(\Omega),\\
\displaystyle\int_\Omega \phi\operatorname{div}\boldsymbol{u}\,\mathrm{d}x = 0 \quad \forall\phi \in L^{\tau'}(\Omega)/\mathbb{R}.
\end{cases}$$

Proof. The proof is based on a density argument and an appropriate splitting of the system. First, we note that $s' < 2 < \varrho$. For each $\varepsilon > 0$, let $\boldsymbol{f}_\varepsilon \in \boldsymbol{H}^{-1}(\Omega)$ and $h_\varepsilon \in H^{-1}(\Omega)$ be such that $\|\boldsymbol{f}_\varepsilon - \boldsymbol{f}\|_{\boldsymbol{W}^{-1,s/2}(\Omega)} < \varepsilon$ and $\|h_\varepsilon - h\|_{W^{-1,s\varrho/(s+\varrho)}(\Omega)} < \varepsilon$. We decompose the problem as the sum of a solution to the following boundary value problem with small data

$$\begin{cases} -\nu\Delta \boldsymbol{u}_{1\varepsilon} + \operatorname{div}(\boldsymbol{u}_{1\varepsilon}\otimes\boldsymbol{u}_{1\varepsilon}) + \nabla P_{1\varepsilon} = \theta_{1\varepsilon}\boldsymbol{g} + \boldsymbol{f} - \boldsymbol{f}_\varepsilon & \text{in } \Omega,\\ \operatorname{div}\, \boldsymbol{u}_{1\varepsilon} = 0 & \text{in } \Omega,\\ -\kappa\Delta\theta_{1\varepsilon} + \operatorname{div}(\theta_{1\varepsilon}\boldsymbol{u}_{1\varepsilon}) = h - h_\varepsilon & \text{in } \Omega,\\ \boldsymbol{u}_{1\varepsilon} = 0, \quad \theta_{1\varepsilon} = 0 & \text{on } \Gamma, \end{cases}$$

and a solution of the perturbed Boussinesq system

$$\begin{cases} -\nu\Delta \boldsymbol{u}_{2\varepsilon} + \operatorname{div}(\boldsymbol{u}_{2\varepsilon}\otimes\boldsymbol{u}_{2\varepsilon} + \boldsymbol{u}_{2\varepsilon}\otimes\boldsymbol{u}_{1\varepsilon} + \boldsymbol{u}_{1\varepsilon}\otimes\boldsymbol{u}_{2\varepsilon}) + \nabla P_{2\varepsilon} = \theta_{2\varepsilon}\boldsymbol{g} + \boldsymbol{f}_\varepsilon & \text{in } \Omega,\\ \operatorname{div} \boldsymbol{u}_{2\varepsilon} = 0 & \text{in } \Omega,\\ -\kappa\Delta\theta_{2\varepsilon} + \operatorname{div}(\theta_{2\varepsilon}\boldsymbol{u}_{2\varepsilon} + \theta_{2\varepsilon}\boldsymbol{u}_{1\varepsilon} + \theta_{1\varepsilon}\boldsymbol{u}_{2\varepsilon}) = h_\varepsilon & \text{in } \Omega,\\ \boldsymbol{u}_{2\varepsilon} = 0, \quad \theta_{2\varepsilon} = 0 & \text{on } \Gamma. \end{cases}$$

Choose $\varepsilon > 0$ small enough so that $(\frac{1}{\nu\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)} + \frac{1}{\nu})\varepsilon < \delta_2$, where $\delta_2 > 0$ is the constant in the statement of Lemma 2.9. Then we have a very weak solution $(\boldsymbol{u}_{1\varepsilon}, \theta_{1\varepsilon}) \in \boldsymbol{L}^s(\Omega) \times L^\varrho(\Omega)$ according to Lemma 2.8. On the other hand, Lemma 2.9 guarantees the existence of a weak solution $(\boldsymbol{u}_{2\varepsilon}, \theta_{2\varepsilon}) \in \boldsymbol{V}^2(\Omega) \times H^1_0(\Omega)$ to the above perturbed Boussinesq system. By direct calculation, one can see that the sum $(\boldsymbol{u}, \theta) = (\boldsymbol{u}_{1\varepsilon} + \boldsymbol{u}_{2\varepsilon}, \theta_{1\varepsilon} + \theta_{2\varepsilon}) \in \boldsymbol{L}^s(\Omega) \times L^\varrho(\Omega) \subset \boldsymbol{L}^{2\tau}(\Omega) \times L^{s\varrho/(s+\varrho)}(\Omega)$ is a very weak solution of the system (2.1). The regularity of the very weak solution follows from Theorem 2.5 and Theorem 2.6 since $\theta\boldsymbol{g} + \boldsymbol{f} - \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) \in \boldsymbol{W}^{-1,\tau}(\Omega)$ and $h - \operatorname{div}(\theta\boldsymbol{u}) \in W^{-1,s\varrho/(s+\varrho)}(\Omega)$, while the existence of a pressure $P \in L^\tau(\Omega)/\mathbb{R}$ follows from de Rham's theorem. ☐

The following result is concerned with the regularity of (very) weak solutions with a restriction to the integrability exponents. As pointed out in [8] for the case of the stationary Navier–Stokes equation, the crucial part for the existence of optimal controls is the linearity of the $\boldsymbol{L}^4$-norm of $\boldsymbol{u}$ appearing on the right-hand side.

Theorem 2.11. *Let $4/3 \le p < 2$ and $2p/(p+1) < q < 2$. Assume that $\boldsymbol{f} \in \boldsymbol{W}^{-1,p}(\Omega)$, $h \in W^{-1,q}(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{L}^\infty(\Omega)$. Then the very weak solution of (2.1) constructed in Theorem 2.10 satisfies $(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W^{1,q}_0(\Omega)$ and the variational equations*

$$\begin{cases} \displaystyle\int_\Omega \{\nu\nabla\boldsymbol{u}\cdot\nabla\boldsymbol{\varphi} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\boldsymbol{\varphi}\}\,\mathrm{d}x - \int_\Omega P \operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x = \int_\Omega \theta\boldsymbol{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}x + \langle \boldsymbol{f}, \boldsymbol{\varphi}\rangle \quad \forall \boldsymbol{\varphi} \in \boldsymbol{W}^{1,p'}_0(\Omega),\\ \displaystyle\int_\Omega \{\kappa\nabla\theta\cdot\nabla\psi + (\boldsymbol{u}\cdot\nabla)\theta\psi\}\,\mathrm{d}x = \langle h, \psi\rangle \quad \forall\psi \in W^{1,q'}_0(\Omega),\\ \displaystyle\int_\Omega \phi \operatorname{div}\boldsymbol{u}\,\mathrm{d}x = 0 \quad \forall \phi \in L^{p'}(\Omega)/\mathbb{R}, \end{cases}$$

hold. Moreover, if $4/3 \le q < 2$ then any weak solution satisfies

$$\begin{aligned} &\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|P\|_{L^p(\Omega)/\mathbb{R}} + \|\theta\|_{W^{1,q}_0(\Omega)} \\ &\qquad \le c(1 + \|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)})(1 + \|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|h\|_{W^{-1,q}(\Omega)})(1 + \|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}) \end{aligned} \tag{2.22}$$

for some constant $c = c_{p,q,\nu,\kappa,\Omega} > 0$ monotonically decreasing in ν and κ.

Proof. Let us take $s = 2p$ and $\varrho = 2pq/(2p-q)$ so that $s\varrho/(s+\varrho) = q$. Then the assumptions on the exponents p and q imply that $8/3 \le s < 4$ and $\varrho > 2$. Therefore, from Theorem 2.10 with $\tau = \min\{\varrho, s/2\} = p$, there exists a weak solution to (2.1) that satisfies

$$(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega).$$

Now, we demonstrate the estimate (2.22). Note that the assumption on p implies that $8/7 \leq 2p/(p+1) < 4/3$. Let $\varphi \in W_0^{1,q}(\Omega)$ be the solution of

$$\begin{cases} -\kappa\Delta\varphi = h & \text{in } \Omega, \\ \varphi = 0 & \text{on } \Gamma, \end{cases}$$

which satisfies the a priori estimate

$$\|\varphi\|_{W_0^{1,q}(\Omega)} \leq \frac{c_P}{\kappa}\|h\|_{W^{-1,q}(\Omega)} \tag{2.23}$$

according to Theorem 2.6. Similarly, let $(\boldsymbol{\phi}, \rho) \in \boldsymbol{V}^p(\Omega) \times L^p(\Omega)/\mathbb{R}$ be the solution of

$$\begin{cases} -\nu\Delta\boldsymbol{\phi} + \nabla\rho = \varphi\boldsymbol{g} + \boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{\phi} = 0 & \text{in } \Omega, \\ \boldsymbol{\phi} = 0 & \text{on } \Gamma, \end{cases}$$

so that by Theorem 2.5, we have the a priori estimate

$$\|\boldsymbol{\phi}\|_{\boldsymbol{V}^p(\Omega)} + \|\rho\|_{L^p(\Omega)/\mathbb{R}} \leq \frac{c_S}{\nu}\left(\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \frac{1}{\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|h\|_{W^{-1,q}(\Omega)}\right), \tag{2.24}$$

where we used the embedding $W_0^{1,q}(\Omega) \subset L^4(\Omega) \subset W^{-1,p}(\Omega)$ along with the estimate (2.23).

Let $(\boldsymbol{v}, \varpi, \eta) = (\boldsymbol{u} - \boldsymbol{\phi}, P - \rho, \theta - \varphi)$. The triple $(\boldsymbol{v}, \varpi, \eta)$ is the weak solution of

$$\begin{cases} -\nu\Delta\boldsymbol{v} + \nabla\varpi = \eta\boldsymbol{g} - \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) & \text{in } \Omega, \\ -\kappa\Delta\eta = -\operatorname{div}(\theta\boldsymbol{u}) & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{v} = 0 & \text{in } \Omega, \\ \boldsymbol{v} = 0, \quad \eta = 0 & \text{on } \Gamma. \end{cases} \tag{2.25}$$

Since $\eta\boldsymbol{g} - \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) \in \boldsymbol{W}^{-1,2}(\Omega)$ and $\operatorname{div}(\theta\boldsymbol{u}) \in W^{-1,2}(\Omega)$ it follows that $(\boldsymbol{v}, \varpi, \eta) \in \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega)$. From $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{\phi}$ and the properties of the trilinear form

$$\langle \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}), \boldsymbol{v}\rangle = -\int_\Omega (\boldsymbol{u}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{u}\,\mathrm{d}x = -\int_\Omega (\boldsymbol{u}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{\phi}\,\mathrm{d}x$$

so that by Hölder's inequality and the embedding $\boldsymbol{V}^2(\Omega) \subset \boldsymbol{V}^p(\Omega) \subset \boldsymbol{L}^4(\Omega)$ we have

$$|\langle \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}), \boldsymbol{v}\rangle| \leq c\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{V}^2(\Omega)}\|\boldsymbol{\phi}\|_{\boldsymbol{V}^p(\Omega)}.$$

Using the test function $\boldsymbol{v}$ in (2.25) yields

$$\|\boldsymbol{v}\|_{\boldsymbol{V}^2(\Omega)} \leq \frac{c}{\nu}(\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|\boldsymbol{\phi}\|_{\boldsymbol{V}^p(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\eta\|_{H_0^1(\Omega)}). \tag{2.26}$$

Similarly, from the equation $\theta = \eta + \varphi$, one has

$$\langle \mathrm{div}(\theta \boldsymbol{u}), \eta\rangle = -\int_\Omega (\boldsymbol{u}\cdot\nabla)\eta\theta \,\mathrm{d}x = -\int_\Omega (\boldsymbol{u}\cdot\nabla)\eta\varphi\,\mathrm{d}x$$

and from the convection-diffusion equation in (2.25) and (2.23), we obtain the estimate

$$\|\eta\|_{H_0^1(\Omega)} \le \frac{c}{\kappa}\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|\varphi\|_{W_0^{1,q}(\Omega)} \le \frac{c}{\kappa^2}\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|h\|_{W^{-1,q}(\Omega)}. \tag{2.27}$$

Combining the estimates for $\boldsymbol{v}$ and η in (2.26) and (2.27), we get

$$\begin{aligned}\|\boldsymbol{v}\|_{\boldsymbol{V}^2(\Omega)} + \|\eta\|_{H_0^1(\Omega)} &\le \frac{c}{\nu}\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|\boldsymbol{\phi}\|_{\boldsymbol{V}^p(\Omega)} + \left(\frac{c}{\nu}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}+1\right)\|\eta\|_{H_0^1(\Omega)}\\ &\le \frac{c}{\nu^2}\left(\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \frac{1}{\kappa}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|h\|_{W^{-1,q}(\Omega)}\right)\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\\ &\quad + \frac{c}{\kappa^2}\left(\frac{c}{\nu}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}+1\right)\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|h\|_{W^{-1,q}(\Omega)}.\end{aligned}$$

Applying the triangle inequality along with the continuous embedding $\boldsymbol{V}^2(\Omega)\times H_0^1(\Omega)\subset \boldsymbol{V}^p(\Omega)\times W_0^{1,q}(\Omega)$ and inequalities (2.23) and (2.25), we obtain the desired estimate (2.22). Furthermore, the a priori estimate for the pressure follows from de Rham's theorem. □

In particular, for measure-valued source terms, the above theorem yields the following.

Corollary 2.12. *Suppose that $p,q\in[4/3,2)$. Let $\boldsymbol{f}_d\in \boldsymbol{W}^{-1,p}(\Omega)$, $h_d\in W^{-1,q}(\Omega)$ and $\boldsymbol{g}\in\boldsymbol{L}^\infty(\Omega)$. Then for every $\boldsymbol{\mu}\in\boldsymbol{M}(\omega_{\mathrm{f}})$ and $\vartheta\in M(\omega_{\mathrm{t}})$, the system (1.2) has a weak solution $(\boldsymbol{u},P,\theta)\in\boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega)$ and we have*

$$\begin{aligned}&\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|P\|_{L^p(\Omega)/\mathbb{R}} + \|\theta\|_{W_0^{1,q}(\Omega)}\\ &\qquad \le c(1+\|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_{\mathrm{f}})} + \|\vartheta\|_{M(\omega_{\mathrm{t}})})(1+\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)})\end{aligned}$$

where $c = c(p,q,\nu,\kappa,\Omega,\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)},\|\boldsymbol{f}_d\|_{\boldsymbol{W}^{-1,p}(\Omega)},\|h_d\|_{W^{-1,q}(\Omega)})>0$ is monotonically decreasing in ν and κ.

For the remaining part of the paper, we shall assume that $p,q\in(4/3,2)$, $\boldsymbol{u}_d\in\boldsymbol{L}^2(\Omega)$, $\theta_d\in L^2(\Omega)$, $\boldsymbol{f}_d\in \boldsymbol{W}^{-1,p}(\Omega)$ and $h_d\in W^{-1,q}(\Omega)$ are fixed.

2.3. Local differentiability at regular points

The possible non-uniqueness of weak solutions to (1.2) makes the analysis of the optimal control problem very difficult to handle. However, for weak solutions that induce unique solutions to the associated linearized state equation, the analysis can be carried out as presented in [8]. Following their approach, we shall refer to a triple $(\boldsymbol{u},P,\theta)\in\boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega)$ that corresponds to a weak solution of (1.2) for a given control $(\boldsymbol{\mu},\vartheta)\in\boldsymbol{M}(\omega_{\mathrm{f}})\times M(\omega_{\mathrm{t}})$ a *regular point* if the linear operator

$$A_{(\boldsymbol{u},P,\theta)}:\boldsymbol{V}^2(\Omega)\times(L^2(\Omega)/\mathbb{R})\times H_0^1(\Omega)\to\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)$$

defined by

$$A_{(\boldsymbol{u},P,\theta)}(\boldsymbol{v},\varpi,\eta) = \begin{pmatrix}-\nu\Delta\boldsymbol{v} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{u} + \nabla\varpi - \eta\boldsymbol{g}\\ -\kappa\Delta\eta + (\boldsymbol{v}\cdot\nabla)\theta + (\boldsymbol{u}\cdot\nabla)\eta\end{pmatrix}$$

is an isomorphism. If the point $(\boldsymbol{u}, P, \theta)$ is clear in the context, we shall use the notation A in place of $A_{(\boldsymbol{u},P,\theta)}$. This regularity assumption is guaranteed as long as the kinematic viscosity is sufficiently large enough as shown in the following theorem. To facilitate the proof, let us denote by $(\boldsymbol{u}_0, P_0, \theta_0)$ a solution corresponding to the null control $(\boldsymbol{\mu}, \vartheta) = (\boldsymbol{0}, 0)$. Observe that $(\boldsymbol{u}_0, P_0, \theta_0, \boldsymbol{0}, 0)$ is a feasible point of the optimal control problem (1.1)–(1.3).

Theorem 2.13. *There exists $\nu_0 > 0$ such that if $\nu > \nu_0$ and $J(\boldsymbol{u}, \theta, \boldsymbol{\mu}, \vartheta) \le J(\boldsymbol{u}_0, \theta_0, \boldsymbol{0}, 0)$ then $(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ is a regular point.*

Proof. Let $j_0 = J(\boldsymbol{u}_0, \theta_0, \boldsymbol{0}, 0)$. From the assumption $J(\boldsymbol{u}, \theta, \boldsymbol{\mu}, \vartheta) \le J(\boldsymbol{u}_0, \theta_0, \boldsymbol{0}, 0)$ we have

$$2\alpha\|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_{\mathrm{f}})} + 2\beta\|\vartheta\|_{M(\omega_{\mathrm{t}})} \le \|\boldsymbol{u}_0 - \boldsymbol{u}_d\|^2_{\boldsymbol{L}^2(\Omega)} + \|\theta_0 - \theta_d\|^2_{\boldsymbol{L}^2(\Omega)}.$$

Fix $\nu_1 > 0$. It follows from Corollary 2.12 that there is a constant $c > 0$ such that for all $\nu \ge \nu_1$ we have

$$\begin{aligned}\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|\theta\|_{W_0^{1,p}(\Omega)} &\le c(1 + \|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}) \\ &\le c_\Omega(1 + \|\boldsymbol{u} - \boldsymbol{u}_d\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{u}_d\|_{\boldsymbol{L}^2(\Omega)}) \le c_\Omega(1 + \sqrt{j_0} + \|\boldsymbol{u}_d\|_{L^2(\Omega)}).\end{aligned}$$

In particular, $(\boldsymbol{u}, \theta)$ is uniformly bounded in $\boldsymbol{L}^4(\Omega) \times L^4(\Omega)$ for $\nu \ge \nu_1$ thanks to the continuous embedding $\boldsymbol{V}^p(\Omega) \times W_0^{1,p}(\Omega) \subset \boldsymbol{L}^4(\Omega) \times L^4(\Omega)$. Let $c_1 > 0$ be such that $\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)} \le c_1$ and $\|\theta\|_{L^4(\Omega)} \le c_1$ for $\nu \ge \nu_1$.

We show that $A : \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega) \to \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega)$ is an isomorphism for sufficiently large ν. To see this, it is enough to establish the coercivity of the bilinear form associated with the operator A. For each $(\boldsymbol{v}, \eta) \in \boldsymbol{V}^2(\Omega) \times H_0^1(\Omega)$, we have

$$\begin{aligned}&\int_\Omega [\nu|\nabla \boldsymbol{v}|^2 + (\boldsymbol{v}\cdot\nabla)\boldsymbol{u}\cdot\boldsymbol{v} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{v} - \eta\boldsymbol{g}\cdot\boldsymbol{v} + \kappa|\nabla\eta|^2 + (\boldsymbol{v}\cdot\nabla)\theta\eta + (\boldsymbol{u}\cdot\nabla)\eta\eta]\,\mathrm{d}x \\ &\quad= \int_\Omega [\nu|\nabla \boldsymbol{v}|^2 - (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{u} - \eta\boldsymbol{g}\cdot\boldsymbol{v} + \kappa|\nabla\eta|^2 - (\boldsymbol{v}\cdot\nabla)\eta\theta]\,\mathrm{d}x.\end{aligned}$$

Denote by c_2, c_3 and c_4 the norms of the continuous embeddings $\boldsymbol{V}^2(\Omega) \subset \boldsymbol{L}^4(\Omega)$, $\boldsymbol{V}^2(\Omega) \subset \boldsymbol{L}^2(\Omega)$ and $H_0^1(\Omega) \subset L^2(\Omega)$, respectively. By Hölder and Young inequalities, we obtain

$$\begin{aligned}-\int_\Omega (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{u}\,\mathrm{d}x &\ge -\|\boldsymbol{v}\|_{\boldsymbol{L}^4(\Omega)}\|\nabla\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)^2}\|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)} \ge -c_1c_2\|\nabla\boldsymbol{v}\|^2_{\boldsymbol{L}^2(\Omega)^2} \\ -\int_\Omega \eta\boldsymbol{g}\cdot\boldsymbol{v} &\ge -\|\eta\|_{L^2(\Omega)}\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)} \ge -c_3c_4\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\nabla\eta\|_{\boldsymbol{L}^2(\Omega)}\|\nabla\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)^2} \\ &\ge -\frac{\kappa}{4}\|\nabla\eta\|^2_{\boldsymbol{L}^2(\Omega)} - \frac{(c_3c_4)^2\|\boldsymbol{g}\|^2_{\boldsymbol{L}^\infty(\Omega)}}{\kappa}\|\nabla\boldsymbol{v}\|^2_{\boldsymbol{L}^2(\Omega)^2} \\ -\int (\boldsymbol{v}\cdot\nabla)\eta\theta\,\mathrm{d}x &\ge -\|\boldsymbol{v}\|_{\boldsymbol{L}^4(\Omega)}\|\nabla\eta\|_{\boldsymbol{L}^2(\Omega)}\|\theta\|_{L^4(\Omega)} \ge -c_1c_2\|\nabla\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)^2}\|\nabla\eta\|_{\boldsymbol{L}^2(\Omega)} \\ &\ge -\frac{\kappa}{4}\|\nabla\eta\|^2_{\boldsymbol{L}^2(\Omega)} - \frac{(c_1c_2)^2}{\kappa}\|\nabla\boldsymbol{v}\|^2_{\boldsymbol{L}^2(\Omega)^2}.\end{aligned}$$

Setting $\nu_2 := c_1c_2 + \kappa^{-1}(c_3c_4)^2\|\boldsymbol{g}\|^2_{\boldsymbol{L}^\infty(\Omega)} + \kappa^{-1}(c_1c_2)^2$ leads to the estimate

$$\int_\Omega [\nu|\nabla \boldsymbol{v}|^2 - (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{u} - \eta\boldsymbol{g}\cdot\boldsymbol{v} + \kappa|\nabla\eta|^2 - (\boldsymbol{v}\cdot\nabla)\eta\theta]\,\mathrm{d}x \ge (\nu - \nu_2)\,\|\nabla\boldsymbol{v}\|^2_{\boldsymbol{L}^2(\Omega)^2} + \frac{\kappa}{2}\|\nabla\eta\|^2_{\boldsymbol{L}^2(\Omega)}.$$

Thus, if $\nu > \nu_0 := \max\{\nu_1, \nu_2\}$ then the bilinear form corresponding to A is coercive. Applying the Lax–Milgram Lemma and invoking de Rham's Theorem for the associated pressure, we see that A is indeed an isomorphism. Therefore, $(\boldsymbol{u}, P, \theta)$ is a regular point. □

For regular points $(\boldsymbol{u}, P, \theta)$, we claim that the corresponding map A can be extended to an isomorphism from $\boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ onto $\boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,p}(\Omega)$. To show this, let us consider the *dual* operator

$$A^\star = A^\star_{(\boldsymbol{u},P,\theta)} : \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega) \to \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega)$$

given by

$$A^\star(\boldsymbol{\varphi}, \pi, \zeta) = \begin{pmatrix} -\nu\Delta\boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi} + (\nabla\boldsymbol{u})^T\boldsymbol{\varphi} + \nabla\pi - \theta\nabla\zeta \\ -\kappa\Delta\zeta - (\boldsymbol{u}\cdot\nabla)\zeta - \boldsymbol{g}\cdot\boldsymbol{\varphi} \end{pmatrix}. \tag{2.28}$$

A simple application of Green's identity and divergence theorem yields the following

$$\langle A(\boldsymbol{v}, \varpi, \eta), (\boldsymbol{\varphi}, \zeta)\rangle = \langle A^\star(\boldsymbol{\varphi}, \pi, \zeta), (\boldsymbol{v}, \eta)\rangle \tag{2.29}$$

for every $(\boldsymbol{v}, \varpi, \eta), (\boldsymbol{\varphi}, \pi, \zeta) \in \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega)$. Indeed, using the anti-symmetry of the trilinear forms associated with the convection terms with respect to the second and third components, we obtain

$$\begin{aligned}
&\langle A(\boldsymbol{v}, \varpi, \eta), (\boldsymbol{\varphi}, \zeta)\rangle \\
&= \int_\Omega \{\nu\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{\varphi} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{v}\cdot\boldsymbol{\varphi} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{u}\cdot\boldsymbol{\varphi}\}\,\mathrm{d}x - \int_\Omega \varpi\,\mathrm{div}\,\boldsymbol{\varphi}\,\mathrm{d}x \\
&\quad - \int_\Omega \eta\boldsymbol{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}x + \int_\Omega \{\kappa\nabla\zeta\cdot\nabla\eta + (\boldsymbol{u}\cdot\nabla)\eta\zeta + (\boldsymbol{v}\cdot\nabla)\theta\zeta\}\,\mathrm{d}x \\
&= \int_\Omega \{\nu\nabla\boldsymbol{\varphi}\cdot\nabla\boldsymbol{v} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi}\cdot\boldsymbol{v} + (\nabla\boldsymbol{u})^T\boldsymbol{\varphi}\cdot\boldsymbol{v} - \theta\nabla\zeta\cdot\boldsymbol{v}\}\,\mathrm{d}x \\
&\quad - \int_\Omega \pi\,\mathrm{div}\,\boldsymbol{v}\,\mathrm{d}x + \int_\Omega \{\kappa\nabla\zeta\cdot\nabla\eta - (\boldsymbol{u}\cdot\nabla)\zeta\eta - (\boldsymbol{g}\cdot\boldsymbol{\varphi})\eta\}\,\mathrm{d}x \\
&= \langle A^\star(\boldsymbol{\varphi}, \pi, \zeta), (\boldsymbol{v}, \eta)\rangle.
\end{aligned}$$

The aforementioned claim on A will be established by a duality argument. In this direction, we shall prove two auxiliary lemmas.

Lemma 2.14. *If $(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ is a regular point then $A^\star : \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega) \to \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega)$ defined by (2.28) is an isomorphism.*

Proof. Take an arbitrary $(\boldsymbol{f}, h) \in \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega)$. Due to the fact that $A = A_{(\boldsymbol{u},P,\theta)}$ is an isomorphism, there is $(\boldsymbol{v}, \varpi, \eta) \in \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega)$ such that $A(\boldsymbol{v}, \varpi, \eta) = (\boldsymbol{f}, h)$, and from (2.29) it holds that

$$\begin{aligned}
|\langle(\boldsymbol{f}, h), (\boldsymbol{\varphi}, \zeta)\rangle| &= |\langle A^\star(\boldsymbol{\varphi}, \pi, \zeta), (\boldsymbol{v}, \eta)\rangle| \\
&\le \|A^\star(\boldsymbol{\varphi}, \pi, \zeta)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)} \|(\boldsymbol{v}, \eta)\|_{\boldsymbol{H}_0^1(\Omega)\times H_0^1(\Omega)} \\
&\le c\|A^\star(\boldsymbol{\varphi}, \pi, \zeta)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)} \|(\boldsymbol{f}, h)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)}.
\end{aligned}$$

By duality, the above inequality implies that

$$\|(\boldsymbol{\varphi}, \zeta)\|_{\boldsymbol{H}_0^1(\Omega)\times H_0^1(\Omega)} \le C\|A^\star(\boldsymbol{\varphi}, \pi, \zeta)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)}. \tag{2.30}$$

It remains to estimate the norm of the *adjoint pressure* π. Let $\boldsymbol{\psi} \in \boldsymbol{H}_0^1(\Omega)$. From the continuous embeddings $\boldsymbol{H}_0^1(\Omega) \subset \boldsymbol{L}^{2p'}(\Omega) \subset \boldsymbol{L}^4(\Omega)$, we obtain the following estimates

$$
\begin{aligned}
|\langle \theta\nabla\zeta, \boldsymbol{\psi}\rangle| &\leq \|\theta\|_{\boldsymbol{L}^4(\Omega)}\|\nabla\zeta\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{L}^4(\Omega)} \leq c\|\theta\|_{W_0^{1,q}(\Omega)}\|\zeta\|_{H_0^1(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{H}_0^1(\Omega)}\\
|\langle (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi}, \boldsymbol{\psi}\rangle| &\leq \|\boldsymbol{u}\|_{\boldsymbol{L}^4(\Omega)}\|\nabla\boldsymbol{\varphi}\|_{\boldsymbol{L}^2(\Omega)^2}\|\boldsymbol{\psi}\|_{\boldsymbol{L}^4(\Omega)} \leq c\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)}\|\boldsymbol{\varphi}\|_{\boldsymbol{H}_0^1(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{H}_0^1(\Omega)}\\
\left|\langle (\nabla\boldsymbol{u})^T\boldsymbol{\varphi}, \boldsymbol{\psi}\rangle\right| &\leq \|\nabla\boldsymbol{u}\|_{\boldsymbol{L}^p(\Omega)^2}\|\boldsymbol{\varphi}\|_{\boldsymbol{L}^{2p'}(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{L}^{2p'}(\Omega)} \leq c\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)}\|\boldsymbol{\varphi}\|_{\boldsymbol{H}_0^1(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{H}_0^1(\Omega)}.
\end{aligned}
$$

Similarly, for $\phi \in H_0^1(\Omega) \subset L^{2q'}(\Omega) \subset L^4(\Omega)$ we deduce that

$$
\begin{aligned}
|\langle \boldsymbol{g}\cdot\boldsymbol{\varphi}, \phi\rangle| &\leq c\|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)}\|\boldsymbol{\varphi}\|_{\boldsymbol{H}_0^1(\Omega)}\|\phi\|_{H_0^1(\Omega)}\\
|\langle (\boldsymbol{u}\cdot\nabla)\zeta, \phi\rangle| &\leq c\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)}\|\zeta\|_{H_0^1(\Omega)}\|\phi\|_{H_0^1(\Omega)}.
\end{aligned}
$$

From these inequalities we immediately obtain that

$$
\begin{aligned}
\|\nabla\pi\|_{\boldsymbol{H}^{-1}(\Omega)} &\leq \|A^\star(\boldsymbol{\varphi},\pi,\zeta)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)}\\
&\quad + c(1+\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|\theta\|_{W_0^{1,q}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^\infty(\Omega)})(\|\boldsymbol{\varphi}\|_{\boldsymbol{H}_0^1(\Omega)} + \|\zeta\|_{H_0^1(\Omega)}),
\end{aligned}
$$

and hence for some constant $c > 0$ independent on $\boldsymbol{\varphi}$, π and ζ, we have from (2.30) that

$$
\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)} + \|\pi\|_{L^2(\Omega)/\mathbb{R}} + \|\zeta\|_{H_0^1(\Omega)} \leq c\|A^\star(\boldsymbol{\varphi},\pi,\zeta)\|_{\boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)}.
$$

In particular, this implies that $A^\star$ is injective and it has a closed range.

If $A^\star$ is not surjective, then for some nonzero $(\boldsymbol{v},\eta) \in \boldsymbol{H}_0^1(\Omega)\times H_0^1(\Omega)$ we would have $\langle A^\star(\boldsymbol{\varphi},\phi,\zeta),(\boldsymbol{v},\eta)\rangle = 0$ for every $(\boldsymbol{\varphi},\phi,\zeta) \in \boldsymbol{V}^2(\Omega)\times(L^2(\Omega)/\mathbb{R})\times H_0^1(\Omega)$. This yields that $\boldsymbol{v} \in \boldsymbol{V}^2(\Omega)$. By de Rham's theorem we have $A(\boldsymbol{v},\varpi,\eta) = 0$ for some $\varpi \in L^2(\Omega)/\mathbb{R}$, and therefore $(\boldsymbol{v},\varpi,\eta) = (\boldsymbol{0},0,0)$ since A is an isomorphism, which is a contradiction, and thus $A^\star$ must be surjective. Therefore, $A^\star$ is an isomorphism. ☐

Lemma 2.15. *If $(\boldsymbol{u},P,\theta) \in \boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega)$ is a regular point then $A^\star$ is an isomorphism from $\boldsymbol{V}^{p'}(\Omega)\times(L^{p'}(\Omega)/\mathbb{R})\times W_0^{1,q'}(\Omega)$ onto $\boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega)$.*

Proof. Since $p' > 2$ and $q' > 2$, we have the continuous embedding $\boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega) \subset \boldsymbol{H}^{-1}(\Omega)\times H^{-1}(\Omega)$. From the previous lemma, for each $(\boldsymbol{f},h) \in \boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega)$ there is a unique $(\boldsymbol{\varphi},\pi,\zeta) \in \boldsymbol{V}^2(\Omega)\times(L^2(\Omega)/\mathbb{R})\times H_0^1(\Omega)$ with $A^\star(\boldsymbol{\varphi},\pi,\zeta) = (\boldsymbol{f},h)$ and

$$
\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)} + \|\pi\|_{L^2(\Omega)/\mathbb{R}} + \|\zeta\|_{H_0^1(\Omega)} \leq c\|(\boldsymbol{f},h)\|_{\boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega)}. \tag{2.31}
$$

We claim that $(\nabla\boldsymbol{u})^T\boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi} - \theta\nabla\zeta \in \boldsymbol{W}^{-1,p'}(\Omega)$. Indeed, take an arbitrary $\boldsymbol{\psi} \in \boldsymbol{W}_0^{1,p}(\Omega)$. From Hölder's inequality and the continuous embeddings $\boldsymbol{W}_0^{1,p}(\Omega) \subset \boldsymbol{L}^{2p/(p-2)}(\Omega)$ and $\boldsymbol{W}_0^{1,2}(\Omega) \subset \boldsymbol{L}^{2p/(3p-4)}(\Omega)$, we obtain

$$
\begin{aligned}
|\langle (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi}, \boldsymbol{\psi}\rangle| = |\langle (\boldsymbol{u}\cdot\nabla)\boldsymbol{\psi}, \boldsymbol{\varphi}\rangle| &\leq \|\boldsymbol{u}\|_{\boldsymbol{L}^{2p/(2-p)}(\Omega)}\|\nabla\boldsymbol{\psi}\|_{\boldsymbol{L}^p(\Omega)^2}\|\boldsymbol{\varphi}\|_{\boldsymbol{L}^{2p/(3p-4)}(\Omega)}\\
&\leq c\|\boldsymbol{u}\|_{\boldsymbol{W}_0^{1,p}(\Omega)}\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)}.
\end{aligned}
$$

Likewise, a similar procedure yields

$$
\begin{aligned}
\left|\langle (\nabla\boldsymbol{u})^T\boldsymbol{\varphi}, \boldsymbol{\psi}\rangle\right| &\leq \|\nabla\boldsymbol{u}\|_{\boldsymbol{L}^p(\Omega)^2}\|\boldsymbol{\varphi}\|_{\boldsymbol{L}^{2p/(3p-4)}(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{L}^{2p/(2-p)}(\Omega)}\\
&\leq c\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)}\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)},
\end{aligned}
$$

and from $W_0^{1,q}(\Omega) \subset L^4(\Omega)$ and $\boldsymbol{W}_0^{1,p}(\Omega) \subset \boldsymbol{L}^4(\Omega)$, we have

$$|\langle \theta \nabla \zeta, \boldsymbol{\psi} \rangle| \le c\|\theta\|_{L^4(\Omega)} \|\nabla \zeta\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{\psi}\|_{\boldsymbol{L}^4(\Omega)} \le c\|\theta\|_{W_0^{1,q}(\Omega)} \|\zeta\|_{H_0^1(\Omega)} \|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)}.$$

Combining the above estimates proves our claim and furthermore

$$\|(\nabla \boldsymbol{u})^T \boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi} - \theta\nabla\zeta\|_{\boldsymbol{W}^{-1,p'}(\Omega)} \le c(\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)} + \|\zeta\|_{H_0^1(\Omega)})$$

for some constant $c > 0$ depending on the norms of $\boldsymbol{u}$ and θ in $\boldsymbol{V}^p(\Omega)$ and $W_0^{1,q}(\Omega)$, respectively. Analogously, one can show that $(\boldsymbol{u}\cdot\nabla)\zeta + \boldsymbol{g}\cdot\boldsymbol{\varphi} \in W^{-1,q'}(\Omega)$ and

$$\|(\boldsymbol{u}\cdot\nabla)\zeta + \boldsymbol{g}\cdot\boldsymbol{\varphi}\|_{W^{-1,q'}(\Omega)} \le c(\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^2(\Omega)} + \|\zeta\|_{H_0^1(\Omega)}).$$

From these observations along with the L^p-theory for the Stokes and Poisson equations, see Theorem 2.5 and Theorem 2.6, we have $(\boldsymbol{\varphi}, \pi, \zeta) \in \boldsymbol{V}^{p'}(\Omega) \times (L^{p'}(\Omega)/\mathbb{R}) \times W_0^{1,q'}(\Omega)$, and moreover from (2.31), we deduce that

$$\|\boldsymbol{\varphi}\|_{\boldsymbol{V}^{p'}(\Omega)} + \|\pi\|_{L^{p'}(\Omega)/\mathbb{R}} + \|\zeta\|_{W_0^{1,q'}(\Omega)} \le c\|(\boldsymbol{f}, h)\|_{\boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega)}.$$

This completes the proof of the lemma. □

Theorem 2.16. *Let $(\boldsymbol{u}, P, \theta) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ be a regular point. Then $A : \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \to \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$ is an isomorphism.*

Proof. We shall proceed by density and duality arguments. Given $(\boldsymbol{f}, h) \in \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$, take a sequence $(\boldsymbol{f}_n, h_n) \in \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega)$ such that $(\boldsymbol{f}_n, h_n) \to (\boldsymbol{f}, h)$ in $\boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$. From Lemma 2.14, there is a unique $(\boldsymbol{v}_n, \varpi_n, \eta_n) \in \boldsymbol{V}^2(\Omega) \times (L^2(\Omega)/\mathbb{R}) \times H_0^1(\Omega)$ such that $A(\boldsymbol{v}_n, \varpi_n, \eta_n) = (\boldsymbol{f}_n, h_n)$. Let $(\boldsymbol{\psi}, \phi) \in \boldsymbol{W}^{-1,p'}(\Omega) \times W^{-1,q'}(\Omega)$ be arbitrary and take $(\boldsymbol{\varphi}, \pi, \zeta) \in \boldsymbol{V}^{p'}(\Omega) \times (L^{p'}(\Omega)/\mathbb{R}) \times W_0^{1,q'}(\Omega)$ such that $A^\star(\boldsymbol{\varphi}, \pi, \zeta) = (\boldsymbol{\psi}, \phi)$, whose existence is guaranteed by Lemma 2.15. Then invoking (2.29), we have

$$\begin{aligned}
|\langle (\boldsymbol{\psi}, \phi), (\boldsymbol{v}_n, \eta_n)\rangle| &= |\langle (\boldsymbol{f}_n, h_n), (\boldsymbol{\varphi}, \zeta)\rangle| \\
&\le \|(\boldsymbol{f}_n, h_n)\|_{\boldsymbol{W}^{-1,p}(\Omega)\times W^{-1,q}(\Omega)} \|(\boldsymbol{\varphi}, \zeta)\|_{\boldsymbol{V}^{p'}(\Omega)\times W_0^{1,q'}(\Omega)} \\
&\le c\|(\boldsymbol{f}_n, h_n)\|_{\boldsymbol{W}^{-1,p}(\Omega)\times W^{-1,q}(\Omega)} \|(\boldsymbol{\psi}, \phi)\|_{\boldsymbol{W}^{-1,p'}(\Omega)\times W^{-1,q'}(\Omega)}.
\end{aligned}$$

Thus, up to a subsequence, we have $(\boldsymbol{v}_n, \eta_n) \rightharpoonup (\boldsymbol{v}, \eta)$ in $\boldsymbol{V}^p(\Omega) \times W_0^{1,q}(\Omega)$, and de Rham's theorem implies the existence of the corresponding pressure term $\varpi \in L^p(\Omega)/\mathbb{R}$ such that $A(\boldsymbol{v}, \varpi, \eta) = (\boldsymbol{f}, h)$. Thus, A is surjective. The injectivity of A follows from integration by parts and the surjectivity of $A^\star$ as stipulated in the previous lemma. □

For regular points, the local existence and uniqueness of weak solutions to the state equation is guaranteed.

Theorem 2.17. *Let $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \bar{\boldsymbol{\mu}}, \bar{\vartheta}) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \times \boldsymbol{M}(\omega_{\mathrm{f}}) \times M(\omega_{\mathrm{t}})$ be such that $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ is a regular point.*

(i) There exists an open, bounded and convex set $\mathcal{O}$ in $\boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$ containing $(\boldsymbol{f}_d + \chi_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\mathrm{t}}}\bar{\vartheta})$, a ball $\mathcal{B}$ in $\boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ containing $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ and a C^∞-map $\mathcal{S} : \mathcal{O} \to \mathcal{B}$ such that $\mathcal{S}(\boldsymbol{f}_d + \chi_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\mathrm{t}}}\bar{\vartheta}) = (\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ and for any $(\boldsymbol{f}, h) \in \mathcal{O}$, the triple $(\boldsymbol{u}, P, \theta) = \mathcal{S}(\boldsymbol{f}, h)$ is the unique solution of (2.1) in $\mathcal{B}$.

(ii) There exists an open, bounded and convex set $\mathcal{U}$ in $\boldsymbol{M}(\omega_{\mathrm{f}}) \times M(\omega_{\mathrm{t}})$ containing $(\bar{\boldsymbol{\mu}}, \bar{\vartheta})$, a ball $\mathcal{B}$ in $\boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ containing $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ and a C^∞-map $S : \mathcal{U} \to \mathcal{B}$ such that $S(\bar{\boldsymbol{\mu}}, \bar{\vartheta}) = (\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ and for every $(\boldsymbol{\mu}, \vartheta) \in \mathcal{U}$, the triple $(\boldsymbol{u}, P, \theta) = S(\boldsymbol{\mu}, \vartheta)$ is the only solution of (1.2) in $\mathcal{B}$.

Proof. To show (i), define the nonlinear operator

$$\mathcal{T} : \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \times \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega) \to \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$$

according to

$$\mathcal{T}(\boldsymbol{u}, P, \theta, \boldsymbol{f}, h) = \begin{pmatrix} -\nu\Delta \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla P - \eta \boldsymbol{g} - \boldsymbol{f} \\ -\kappa\Delta\eta + (\boldsymbol{u}\cdot\nabla)\theta - h \end{pmatrix}.$$

Observe that $\mathcal{T}$ is of class C^∞, and moreover $\mathcal{T}(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \boldsymbol{f}_d + \chi_{\omega_{\rm f}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\rm t}}\bar{\vartheta}) = (\boldsymbol{0}, 0)$. Also, the fact that $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ is regular implies that the following operator is an isomorphism

$$A_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})} = \frac{\partial \mathcal{T}}{\partial(\boldsymbol{u}, P, \theta)}(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \boldsymbol{f}_d + \chi_{\omega_{\rm f}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\rm t}}\bar{\vartheta}).$$

Then part (i) is a consequence of the implicit function theorem, see [28, Section 4.7].

To prove (ii), define $\mathcal{I} : \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t}) \to \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$ by $\mathcal{I}(\boldsymbol{\mu}, \vartheta) = (\boldsymbol{f}_d + \chi_{\omega_{\rm f}}\boldsymbol{\mu}, h_d + \chi_{\omega_{\rm t}}\vartheta)$. Since $\mathcal{I}$ is affine, it is of class C^∞. Then one may take $S = \mathcal{S}\circ\mathcal{I}$ and $\mathcal{O}$ such that $\mathcal{I}^{-1}(\mathcal{O}) = \mathcal{U}$. □

We end this section by presenting the first and second-order derivatives of the local control-to-state operator $\mathcal{S}$ given in part (i) of the previous theorem. For $\mathbf{q} = (\boldsymbol{f}, h) \in \mathcal{O}$, we set

$$(\boldsymbol{u}_{\mathbf{q}}, P_{\mathbf{q}}, \theta_{\mathbf{q}}) = \mathcal{S}(\boldsymbol{f}, h) = \mathcal{S}(\mathbf{q}).$$

Given a direction $\mathbf{r} = (\delta\boldsymbol{f}, \delta h) \in \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$, we have $(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}}) = \mathcal{S}'(\boldsymbol{f}, h)(\delta\boldsymbol{f}, \delta h)$ if and only if $A_{(\boldsymbol{u}_{\mathbf{q}},P_{\mathbf{q}},\theta_{\mathbf{q}})}(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}}) = (\delta\boldsymbol{f}, \delta h)$, that is, $(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}})$ is a weak solution of

$$\begin{cases} -\nu\Delta\boldsymbol{v}_{\mathbf{r}} + (\boldsymbol{u}_{\mathbf{q}}\cdot\nabla)\boldsymbol{v}_{\mathbf{r}} + (\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)\boldsymbol{u}_{\mathbf{q}} + \nabla\varpi_{\mathbf{r}} = \eta_{\mathbf{r}}\boldsymbol{g} + \chi_{\omega_{\rm f}}\delta\boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{v}_{\mathbf{r}} = 0 & \text{in } \Omega, \\ -\kappa\Delta\eta_{\mathbf{r}} + (\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)\theta_{\mathbf{q}} + (\boldsymbol{u}_{\mathbf{q}}\cdot\nabla)\eta_{\mathbf{r}} = \chi_{\omega_{\rm t}}\delta h & \text{in } \Omega, \\ \boldsymbol{v}_{\mathbf{r}} = 0, \quad \eta_{\mathbf{r}} = 0 & \text{on } \Gamma. \end{cases}$$

For the second derivative, if $\mathbf{r}_i = (\delta\boldsymbol{f}_i, \delta h_i) \in \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$ for $i = 1, 2$ then

$$(\boldsymbol{v}, \varpi, \eta) = \mathcal{S}''(\boldsymbol{f}, h)((\delta\boldsymbol{f}_1, \delta h_1), (\delta\boldsymbol{f}_2, \delta h_2))$$

if and only if

$$A_{(\boldsymbol{u}_{\mathbf{q}},p_{\mathbf{q}},\theta_{\mathbf{q}})}(\boldsymbol{v}, \varpi, \eta) = -((\boldsymbol{v}_{\mathbf{r}_1}\cdot\nabla)\boldsymbol{v}_{\mathbf{r}_2} + (\boldsymbol{v}_{\mathbf{r}_2}\cdot\nabla)\boldsymbol{v}_{\mathbf{r}_1}, (\boldsymbol{v}_{\mathbf{r}_1}\cdot\nabla)\theta_{\mathbf{r}_2} + (\boldsymbol{v}_{\mathbf{r}_2}\cdot\nabla)\theta_{\mathbf{r}_1})$$

where $(\boldsymbol{v}_{\mathbf{r}_i}, \varpi_{\mathbf{r}_i}, \eta_{\mathbf{r}_i}) = \mathcal{S}'(\boldsymbol{f}, h)(\delta\boldsymbol{f}_i, \delta h_i)$, that is, $(\boldsymbol{v}, \varpi, \eta)$ is a weak solution of

$$\begin{cases} -\nu\Delta\boldsymbol{v} + (\boldsymbol{u}_{\mathbf{q}}\cdot\nabla)\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{u}_{\mathbf{q}} + \nabla\varpi = \eta\boldsymbol{g} - (\boldsymbol{v}_{\mathbf{r}_1}\cdot\nabla)\boldsymbol{v}_{\mathbf{r}_2} - (\boldsymbol{v}_{\mathbf{r}_2}\cdot\nabla)\boldsymbol{v}_{\mathbf{r}_1} & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{v} = 0, & \text{in } \Omega, \\ -\kappa\Delta\eta + (\boldsymbol{v}\cdot\nabla)\theta_{\mathbf{q}} + (\boldsymbol{u}_{\mathbf{q}}\cdot\nabla)\eta = -(\boldsymbol{v}_{\mathbf{r}_1}\cdot\nabla)\theta_{\mathbf{r}_2} - (\boldsymbol{v}_{\mathbf{r}_2}\cdot\nabla)\theta_{\mathbf{r}_1} & \text{in } \Omega, \\ \boldsymbol{v} = 0, \quad \eta = 0 & \text{on } \Gamma. \end{cases}$$

These representations can be verified by the usual differential calculus in Banach spaces. Analogous representations also hold for right-hand sides in the open set $\mathcal{U} \subset \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$ corresponding to the local solution operator S stated in part (ii) of Theorem 2.17.

3. Analysis of the optimal control problem

In this section we prove the existence of optimal controls and characterize the first-order necessary condition for optimality.

Theorem 3.1. *The optimal control problem* (1.1)–(1.3) *has at least one solution*

$$(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \bar{\boldsymbol{\mu}}, \bar{\vartheta}) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \times \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t}).$$

Proof. Recall that $(\boldsymbol{u}_0, P_0, \theta_0)$ is a solution corresponding to the null control $(\boldsymbol{\mu}, \vartheta) = (\boldsymbol{0}, 0)$. Consider a minimizing sequence $\{(\boldsymbol{u}_n, P_n, \theta_n, \boldsymbol{\mu}_n, \vartheta_n)\}_n \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \times \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$ so that

$$J(\boldsymbol{u}_n, \theta_n, \boldsymbol{\mu}_n, \vartheta_n) \le J(\boldsymbol{u}_0, \theta_0, \boldsymbol{0}, 0) = \frac{1}{2}\|\boldsymbol{u}_0 - \boldsymbol{u}_d\|_{\boldsymbol{L}^2(\Omega)} + \frac{1}{2}\|\theta_0 - \theta_d\|_{L^2(\Omega)}$$

and the left-hand side tends to the infimum of J as $n \to \infty$. Thus, the sequence $\{(\boldsymbol{u}_{\boldsymbol{n}}, \theta_n, \boldsymbol{\mu}_n, \vartheta_n)\}_n$ is bounded in $\boldsymbol{L}^2(\Omega) \times L^2(\Omega) \times \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$ by the definition of J. In particular, by the Banach–Alaoglu–Bourbaki Theorem applied to $\boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$ having the separable predual space $\boldsymbol{C}_0(\omega_{\rm f}) \times C_0(\omega_{\rm t})$, up to a subsequence we have $\boldsymbol{\mu}_n \overset{*}{\rightharpoonup} \boldsymbol{\mu}$ in $\boldsymbol{M}(\omega_{\rm f})$ and $\vartheta_n \overset{*}{\rightharpoonup} \vartheta$ in $M(\omega_{\rm t})$. On the other hand, from the stability of weak solutions in Corollary 2.12, one has

$$\|\boldsymbol{u}_n\|_{\boldsymbol{V}^p(\Omega)} + \|P_n\|_{L^p(\Omega)/\mathbb{R}} + \|\theta_n\|_{W_0^{1,q}(\Omega)} \le c(1 + \|\boldsymbol{u}_n\|_{\boldsymbol{L}^4(\Omega)})$$

for some constant $c > 0$ independent on n.

Due to the fact that $2p/(2-p) > 4$, by applying Lion's Lemma to the compact embedding $\boldsymbol{W}_0^{1,p}(\Omega) \subset \boldsymbol{L}^4(\Omega)$ and the continuous embedding $\boldsymbol{L}^4(\Omega) \subset \boldsymbol{L}^2(\Omega)$, we have for every $\varepsilon > 0$ the existence of $c_\varepsilon > 0$ such that

$$\|\boldsymbol{u}_n\|_{\boldsymbol{L}^4(\Omega)} \le \varepsilon\|\boldsymbol{u}_n\|_{\boldsymbol{V}^p(\Omega)} + c_\varepsilon\|\boldsymbol{u}_n\|_{\boldsymbol{L}^2(\Omega)}.$$

Choosing $\varepsilon > 0$ small enough so that $c\varepsilon < 1$ yields the boundedness of $\{\boldsymbol{u}_n\}_n$ in $\boldsymbol{V}^p(\Omega)$, $\{P_n\}_n$ in $L^p(\Omega)/\mathbb{R}$ and $\{\theta_n\}_n$ in $W_0^{1,q}(\Omega)$. Hence, up to a subsequence once more, it holds that $\boldsymbol{u}_n \rightharpoonup \boldsymbol{u}$ in $\boldsymbol{V}^p(\Omega)$, $P_n \rightharpoonup P$ in $L^p(\Omega)/\mathbb{R}$ and $\theta_n \rightharpoonup \theta$ in $W_0^{1,q}(\Omega)$.

We show that $(\boldsymbol{u}, P, \theta)$ is a weak solution corresponding to the control $(\boldsymbol{\mu}, \vartheta)$. This can be verified by passing to the limit in the weak formulation satisfied by the minimizing sequence. The only critical part in the passage of limit is the convergence of the trilinear terms in the fluid and heat equations. In virtue of the compactness of the Sobolev embeddings $\boldsymbol{W}^{1,p}(\Omega) \subset \boldsymbol{L}^{2p}(\Omega)$ and $W^{1,q}(\Omega) \subset L^{2q}(\Omega)$, it follows by further extracting a subsequence that $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\boldsymbol{L}^{2p}(\Omega)$ and $\theta_n \to \theta$ in $L^{2q}(\Omega)$. Hence

$$\begin{aligned}&\left|\int_\Omega (\boldsymbol{u}_n \cdot \nabla)\boldsymbol{\psi} \cdot \boldsymbol{u}_n \,\mathrm{d}x - \int_\Omega (\boldsymbol{u} \cdot \nabla)\boldsymbol{\psi} \cdot \boldsymbol{u} \,\mathrm{d}x\right| \\ &\qquad \le c_\Omega \|\boldsymbol{u}_n - \boldsymbol{u}\|_{\boldsymbol{L}^{2p}(\Omega)} \|\nabla\boldsymbol{\psi}\|_{\boldsymbol{L}^{p'}(\Omega)^2} (\|\boldsymbol{u}_n\|_{\boldsymbol{L}^{2p}(\Omega)} + \|\boldsymbol{u}\|_{\boldsymbol{L}^{2p}(\Omega)}) \to 0\end{aligned}$$

for every $\boldsymbol{\psi} \in \boldsymbol{W}_0^{1,p'}(\Omega)$. On the other hand, from the assumptions on p and q, that is $p, q \in (4/3, 2)$, we have $q < 2 \leq p/(2-p)$, and so

$$\frac{2-q}{2q} + \frac{1}{q'} + \frac{1}{2p} = \frac{1}{2} + \frac{1}{2p} < \frac{7}{8}$$
$$\frac{1}{2q} + \frac{1}{q'} + \frac{2-p}{2p} \leq \frac{1}{2} + \frac{1}{2q'} + \frac{2-p}{2p} \leq \frac{3}{4} + \frac{1}{2q'} < 1.$$

These two inequalities allow us to apply Hölder's inequality to obtain

$$\left| \int_\Omega \theta_n \boldsymbol{u}_n \cdot \nabla\phi \, \mathrm{d}x - \int_\Omega \theta \boldsymbol{u} \cdot \nabla\phi \, \mathrm{d}x \right|$$
$$\leq c_\Omega (\|\theta_n - \theta\|_{L^{2q}(\Omega)} \|\boldsymbol{u}_n\|_{\boldsymbol{L}^{2p/(2-p)}(\Omega)} + \|\theta\|_{L^{2q/(2-q)}(\Omega)} \|\boldsymbol{u}_n - \boldsymbol{u}\|_{\boldsymbol{L}^{2p}(\Omega)}) \|\nabla\phi\|_{\boldsymbol{L}^{q'}(\Omega)} \to 0$$

for every $\phi \in W_0^{1,q'}(\Omega)$.

Thus, $(\boldsymbol{u}, P, \theta)$ is a weak solution associated with the control $(\boldsymbol{\mu}, \vartheta)$. From the lower semicontinuity of the norm with respect to the weak and weak* topologies, it follows that $(\boldsymbol{u}, P, \theta, \boldsymbol{\mu}, \vartheta)$ is a solution of the optimal control problem. □

Let us denote the reduced cost functional $j : \mathcal{U} \to \mathbb{R}$ by

$$j(\boldsymbol{\mu}, \vartheta) = j(\mathbf{q}) = \frac{1}{2}\|\boldsymbol{u}_\mathbf{q} - \boldsymbol{u}_d\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\theta_\mathbf{q} - \theta_d\|_{L^2(\Omega)}^2 + \alpha\|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_\mathrm{f})} + \beta\|\vartheta\|_{M(\omega_\mathrm{t})}$$

where $(\boldsymbol{u}_\mathbf{q}, P_\mathbf{q}, \theta_\mathbf{q}) = S(\boldsymbol{\mu}, \vartheta)$. Decompose j to a smooth part and convex parts as follows

$$j_d(\mathbf{q}) = \frac{1}{2}\|\boldsymbol{u}_\mathbf{q} - \boldsymbol{u}_d\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\theta_\mathbf{q} - \theta_d\|_{L^2(\Omega)}^2, \quad j_\mathrm{f}(\boldsymbol{\mu}) = \|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_\mathrm{f})}, \quad j_\mathrm{t}(\vartheta) = \|\vartheta\|_{M(\omega_\mathrm{t})}.$$

The tracking part j_d of the cost functional j is C^∞ and its first and second directional derivatives are given by, for $\mathbf{r} = (\delta\boldsymbol{\mu}, \delta\vartheta) \in \boldsymbol{M}(\omega_\mathrm{f}) \times M(\omega_\mathrm{t})$

$$j_d'(\mathbf{q})\mathbf{r} = \int_{\omega_\mathrm{f}} \boldsymbol{\varphi}_\mathbf{q} \, \mathrm{d}(\delta\boldsymbol{\mu}) + \int_{\omega_\mathrm{t}} \zeta_\mathbf{q} \, \mathrm{d}(\delta\vartheta) \tag{3.1}$$

$$j_d''(\mathbf{q})(\mathbf{r}, \mathbf{r}) = \int_\Omega (|\boldsymbol{v}_\mathbf{r}|^2 + |\eta_\mathbf{r}|^2 - 2(\boldsymbol{v}_\mathbf{r} \cdot \nabla)\boldsymbol{v}_\mathbf{r} \cdot \boldsymbol{\varphi}_\mathbf{q} - 2(\boldsymbol{v}_\mathbf{r} \cdot \nabla)\eta_\mathbf{r}\zeta_\mathbf{q}) \, \mathrm{d}x \tag{3.2}$$

where $(\boldsymbol{v}_\mathbf{r}, \varpi_\mathbf{r}, \eta_\mathbf{r}) = S'(\mathbf{q})\mathbf{r}$ and $(\boldsymbol{\varphi}_\mathbf{q}, \pi_\mathbf{q}, \zeta_\mathbf{q}) \in \boldsymbol{V}^{p'}(\Omega) \times (L^{p'}(\Omega)/\mathbb{R}) \times W_0^{1,q'}(\Omega)$ is the solution of the linear adjoint system

$$\begin{cases} -\nu\Delta\boldsymbol{\varphi}_\mathbf{q} - (\boldsymbol{u}_\mathbf{q} \cdot \nabla)\boldsymbol{\varphi}_\mathbf{q} + (\nabla\boldsymbol{u}_\mathbf{q})^T\boldsymbol{\varphi}_\mathbf{q} + \nabla\pi_\mathbf{q} = \theta_\mathbf{q}\nabla\zeta_\mathbf{q} + \boldsymbol{u}_\mathbf{q} - \boldsymbol{u}_d & \text{in } \Omega, \\ \operatorname{div}\boldsymbol{\varphi}_\mathbf{q} = 0 & \text{in } \Omega, \\ -\kappa\Delta\zeta_\mathbf{q} - (\boldsymbol{u}_\mathbf{q} \cdot \nabla)\zeta_\mathbf{q} = \boldsymbol{g} \cdot \boldsymbol{\varphi}_\mathbf{q} + \theta_\mathbf{q} - \theta_d & \text{in } \Omega, \\ \boldsymbol{\varphi}_\mathbf{q} = 0, \quad \zeta_\mathbf{q} = 0 & \text{on } \Gamma, \end{cases}$$

or alternatively, in terms of the operator $A^\star$, we have

$$A^\star_{(\boldsymbol{u}_\mathbf{q}, P_\mathbf{q}, \theta_\mathbf{q})}(\boldsymbol{\varphi}_\mathbf{q}, \pi_\mathbf{q}, \zeta_\mathbf{q}) = (\boldsymbol{u}_\mathbf{q} - \boldsymbol{u}_d, \theta_\mathbf{q} - \theta_d). \tag{3.3}$$

On the other hand, due to Lipschitz continuity and convexity, the directional derivatives $j_{\mathrm{f}}'(\boldsymbol{\mu};\delta\boldsymbol{\mu})$ and $j_{\mathrm{t}}'(\vartheta;\delta\vartheta)$ at $\boldsymbol{\mu}\in\boldsymbol{M}(\omega_{\mathrm{f}})$ and $\vartheta\in M(\omega_{\mathrm{t}})$ exist in every directions $\delta\boldsymbol{\mu}\in\boldsymbol{M}(\omega_{\mathrm{f}})$ and $\delta\vartheta\in M(\omega_{\mathrm{t}})$, respectively. Moreover, j_{f} and j_{t} are subdifferentiable. Recall that $\boldsymbol{\lambda}\in\partial j_{\mathrm{f}}(\boldsymbol{\mu})\subset\boldsymbol{M}(\omega_{\mathrm{f}})^*$ if

$$\langle\boldsymbol{\lambda},\delta\boldsymbol{\mu}-\boldsymbol{\mu}\rangle_{\boldsymbol{M}(\omega_{\mathrm{f}})^*\times\boldsymbol{M}(\omega_{\mathrm{f}})}+j_{\mathrm{f}}(\boldsymbol{\mu})\le j_{\mathrm{f}}(\delta\boldsymbol{\mu})\quad\forall\delta\boldsymbol{\mu}\in\boldsymbol{M}(\omega_{\mathrm{f}})$$

or equivalently if $\langle\boldsymbol{\lambda},\boldsymbol{\mu}\rangle_{\boldsymbol{M}(\omega_{\mathrm{f}})^*\times\boldsymbol{M}(\omega_{\mathrm{f}})}=j_{\mathrm{f}}(\boldsymbol{\mu})$ and $\langle\boldsymbol{\lambda},\delta\boldsymbol{\mu}\rangle_{\boldsymbol{M}(\omega_{\mathrm{f}})^*\times\boldsymbol{M}(\omega_{\mathrm{f}})}\le j_{\mathrm{f}}(\delta\boldsymbol{\mu})$ for all $\delta\boldsymbol{\mu}\in\boldsymbol{M}(\omega_{\mathrm{f}})$. Similar remarks are applicable to j_{t}.

Let us turn to the local first-order necessary condition for optimal solutions that are regular. We say that $(\bar{\boldsymbol{u}},\bar{P},\bar{\theta},\bar{\boldsymbol{\mu}},\bar{\vartheta})$ is a *local solution* of the optimal control problem (1.1)–(1.3) if there exists an open set O in $\boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega)\times\boldsymbol{M}(\omega_{\mathrm{f}})\times M(\omega_{\mathrm{t}})$ containing $(\bar{\boldsymbol{u}},\bar{P},\bar{\theta},\bar{\boldsymbol{\mu}},\bar{\vartheta})$ such that $J(\boldsymbol{u},P,\theta,\boldsymbol{\mu},\vartheta)\le J(\bar{\boldsymbol{u}},\bar{P},\bar{\theta},\bar{\boldsymbol{\mu}},\bar{\vartheta})$ for every $(\boldsymbol{u},P,\theta,\boldsymbol{\mu},\vartheta)\in O$.

Theorem 3.2. *Let $(\bar{\boldsymbol{\mu}},\bar{\vartheta})\in\boldsymbol{M}(\omega_{\mathrm{f}})\times M(\omega_{\mathrm{t}})$ be a local solution of (1.1)–(1.3) with corresponding state $(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})\in\boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega)$ that is regular. Then there exists $(\bar{\boldsymbol{\varphi}},\bar{\pi},\bar{\zeta})\in\boldsymbol{V}^{p'}(\Omega)\times(L^{p'}(\Omega)/\mathbb{R})\times W_0^{1,q'}(\Omega)$ such that $-\frac{1}{\alpha}\bar{\boldsymbol{\varphi}}\in\partial j_{\mathrm{f}}(\bar{\boldsymbol{\mu}})\cap\boldsymbol{C}_0(\omega_{\mathrm{f}})$, $-\frac{1}{\beta}\bar{\zeta}\in\partial j_{\mathrm{t}}(\bar{\vartheta})\cap C_0(\omega_{\mathrm{t}})$ and*

$$A^\star_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})}(\bar{\boldsymbol{\varphi}},\bar{\pi},\bar{\zeta})=(\bar{\boldsymbol{u}}-\boldsymbol{u}_d,\bar{\theta}-\theta_d).\tag{3.4}$$

Proof. According to the convexity of j, the local optimality of $(\bar{\boldsymbol{\mu}},\bar{\vartheta})$ and (3.1), we have

$$\begin{aligned}0&\le\lim_{\sigma\downarrow0}\frac{1}{\sigma}\{j(\bar{\boldsymbol{\mu}}+\sigma(\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}),\bar{\vartheta}+\sigma(\vartheta-\bar{\vartheta}))-j(\bar{\boldsymbol{\mu}},\bar{\vartheta})\}\\&\le\lim_{\sigma\downarrow0}\frac{1}{\sigma}\{j_d(\bar{\boldsymbol{\mu}}+\sigma(\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}),\bar{\vartheta}+\sigma(\vartheta-\bar{\vartheta}))-j_d(\bar{\boldsymbol{\mu}},\bar{\vartheta})\}+\alpha(j_{\mathrm{f}}(\boldsymbol{\mu})-j_{\mathrm{f}}(\bar{\boldsymbol{\mu}}))+\beta(j_{\mathrm{t}}(\vartheta)-j_{\mathrm{t}}(\bar{\vartheta}))\\&=\int_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\varphi}}\,\mathrm{d}(\boldsymbol{\mu}-\bar{\boldsymbol{\mu}})+\int_{\omega_{\mathrm{t}}}\bar{\zeta}\,\mathrm{d}(\vartheta-\bar{\vartheta})+\alpha(\|\boldsymbol{\mu}\|_{\boldsymbol{M}(\omega_{\mathrm{f}})}-\|\bar{\boldsymbol{\mu}}\|_{\boldsymbol{M}(\omega_{\mathrm{f}})})+\beta(\|\vartheta\|_{M(\omega_{\mathrm{t}})}-\|\bar{\vartheta}\|_{M(\omega_{\mathrm{t}})})\end{aligned}$$

for all $(\boldsymbol{\mu},\vartheta)\in\boldsymbol{M}(\omega_{\mathrm{f}})\times M(\omega_{\mathrm{t}})$, with $(\bar{\boldsymbol{\varphi}},\bar{\pi},\bar{\zeta})$ given by (3.4). Choosing $\vartheta=\bar{\vartheta}$ yields that $-\frac{1}{\alpha}\bar{\boldsymbol{\varphi}}\in\partial j_{\mathrm{f}}(\bar{\boldsymbol{\mu}})$, while $\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}$ gives us $-\frac{1}{\beta}\bar{\zeta}\in\partial j_{\mathrm{t}}(\bar{\vartheta})$. The remaining part is a consequence of the embeddings $\boldsymbol{W}_0^{1,p'}(\Omega)\subset\boldsymbol{C}_0(\omega_{\mathrm{f}})$ and $W_0^{1,q'}(\Omega)\subset C_0(\omega_{\mathrm{t}})$ since $p',q'>2$. □

The sparsity of measure controls is given in the following theorem, whose proof can be obtained by adapting the methods in Corollary 3.8 of [8], therefore the proof is omitted.

Theorem 3.3. *With the notations of the previous theorem and assuming that $\bar{\mu}_1$, $\bar{\mu}_2$ and $\bar{\vartheta}$ are not equal zero, then $\|\bar{\varphi}_1\|_\infty=\|\bar{\varphi}_2\|_\infty=\alpha$, $\|\bar{\zeta}\|_\infty=\beta$ and*

$$\begin{aligned}&\mathrm{supp}(\bar{\mu}_i^\pm)\subset\{x\in\omega_{\mathrm{f}}:\bar{\varphi}(x)=\mp\alpha\},\\&\mathrm{supp}(\bar{\vartheta}^\pm)\subset\{x\in\omega_{\mathrm{t}}:\bar{\zeta}(x)=\mp\beta\},\end{aligned}$$

where $\bar{\mu}_i=\bar{\mu}_i^+-\bar{\mu}_i^-$ for $i=1,2$ and $\bar{\vartheta}=\bar{\vartheta}^+-\bar{\vartheta}^-$ are the Jordan decompositions of the Borel measures $\bar{\mu}_i$ and $\bar{\vartheta}$, respectively.

Remark 3.4. All the results in this section can be adapted to the case where there is control only in the fluid equation ($\omega_{\mathrm{t}}=\emptyset$ and $\beta=0$) or only in the convection-diffusion equation ($\omega_{\mathrm{f}}=\emptyset$ and $\alpha=0$).

4. Second-order optimality conditions

We now establish the second-order necessary and sufficient optimality conditions for regular local solutions as defined in Section 2.3. All throughout this section, $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \bar{\boldsymbol{\mu}}, \bar{\vartheta})$ will denote a local solution of (1.1)–(1.3) for which the triple $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ is regular. Moreover, let $(\bar{\boldsymbol{\varphi}}, \bar{\pi}, \bar{\zeta})$ be the adjoint state satisfying the first-order optimality condition in Theorem 3.2. We shall also use the notation $\bar{\mathbf{q}} = (\bar{\boldsymbol{\mu}}, \bar{\vartheta})$ for the optimal control.

For $\tau \geq 0$, we define the cone of critical directions

$$\mathcal{C}_\tau(\bar{\boldsymbol{\mu}}, \bar{\vartheta}) = \{\mathbf{q} = (\boldsymbol{\mu}, \vartheta) \in \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t}) : j'(\bar{\boldsymbol{\mu}}, \bar{\vartheta})(\boldsymbol{\mu}, \vartheta) \leq \tau(\|\boldsymbol{v}_{\mathbf{q}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}}\|_{L^2(\Omega)})\}$$

where $(\boldsymbol{v}_{\mathbf{q}}, \varpi_{\mathbf{q}}, \eta_{\mathbf{q}}) = S'(\bar{\mathbf{q}})\mathbf{q}$. The second-order necessary condition is formulated in the cone $\mathcal{C}_0(\bar{\boldsymbol{\mu}}, \bar{\vartheta})$, while the second-order sufficient condition will be developed in the *slightly* larger cone $\mathcal{C}_\tau(\bar{\boldsymbol{\mu}}, \bar{\vartheta})$ for some $\tau > 0$, see [7, 8, 27] for instance.

From the convexity of j, it follows that $\mathcal{C}_\tau(\bar{\boldsymbol{\mu}}, \bar{\vartheta})$ is a convex cone in $\boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$. By employing the methods in [7] through the Lebesgue decomposition of regular Borel measures, the following second-order necessary condition can be established.

Theorem 4.1. *Let $p, q \in (4/3, 2)$. If $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta}, \bar{\boldsymbol{\mu}}, \bar{\vartheta}) \in \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \times \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$ is a local solution of (1.1)–(1.3) such that $(\bar{\boldsymbol{u}}, \bar{P}, \bar{\theta})$ is regular, then we have $j_d''(\bar{\boldsymbol{\mu}}, \bar{\vartheta})(\boldsymbol{\mu}, \vartheta)^2 \geq 0$ for every $(\boldsymbol{\mu}, \vartheta) \in \mathcal{C}_0(\bar{\boldsymbol{\mu}}, \bar{\vartheta})$.*

To formulate the second-order sufficient conditions, we shall restrict the powers p and q so that $4/3 < q \leq p < 2$. This additional assumption on q is imposed in order to successfully estimate the convection term in the linearized heat equation.

Recall from Theorem 2.17 that there is a bounded, open and convex set $\mathcal{O}$ such that $\mathcal{S}'(\boldsymbol{f}, h) : \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega) \to \boldsymbol{V}^p(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega)$ is an isomorphism with uniformly bounded operator norms over all $(\boldsymbol{f}, h) \in \mathcal{O}$, *i.e.* for some $c > 0$ it holds that

$$\sup_{(\boldsymbol{f},h)\in\mathcal{O}} \|\mathcal{S}'(\boldsymbol{f}, h)\|_{\mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega)\times W^{-1,q}(\Omega), \boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega))} \leq c. \tag{4.1}$$

Furthermore, from the proof of Theorem 2.17(ii) we have $\mathcal{U} = \mathcal{I}^{-1}(\mathcal{O}) = (\mathcal{O} - (\boldsymbol{f}_d, h_d)) \cap (\boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t}))$ thanks to the embedding $\boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t}) \subset \boldsymbol{W}^{-1,p}(\Omega) \times W^{-1,q}(\Omega)$. Therefore, we have the local Lipschitz continuity of the solution operator, *i.e.* for some constant $c > 0$ it holds that

$$\begin{aligned} &\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{V}^p(\Omega)} + \|P_{\mathbf{q}} - \bar{P}\|_{L^p(\Omega)/\mathbb{R}} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{W_0^{1,q}(\Omega)} \\ &\qquad \leq c(\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)}), \qquad \forall \mathbf{q} = (\boldsymbol{\mu}, \vartheta) \in \mathcal{U}. \end{aligned} \tag{4.2}$$

As a consequence of (4.1) and (4.2), we have

$$\sup_{\mathbf{q}\in\mathcal{U}} \|A^{-1}_{(\boldsymbol{u}_{\mathbf{q}}, P_{\mathbf{q}}, \theta_{\mathbf{q}})}\|_{\mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega)\times W^{-1,q}(\Omega), \boldsymbol{V}^p(\Omega)\times(L^p(\Omega)/\mathbb{R})\times W_0^{1,q}(\Omega))} \leq c. \tag{4.3}$$

Lemma 4.2. *Let $\mathbf{q} = (\boldsymbol{\mu}, \vartheta) \in \mathcal{U}$ and $\mathbf{r} = (\delta\boldsymbol{\mu}, \delta\vartheta) \in \boldsymbol{M}(\omega_{\rm f}) \times M(\omega_{\rm t})$. Suppose that $(\boldsymbol{v}_{\mathbf{q},\mathbf{r}}, \varpi_{\mathbf{q},\mathbf{r}}, \eta_{\mathbf{q},\mathbf{r}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\rm f}}\boldsymbol{\mu}, h_d + \chi_{\omega_{\rm t}}\vartheta)(\chi_{\omega_{\rm f}}\delta\boldsymbol{\mu}, \chi_{\omega_{\rm t}}\delta\vartheta)$ and $(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\rm f}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\rm t}}\bar{\vartheta})(\chi_{\omega_{\rm f}}\delta\boldsymbol{\mu}, \chi_{\omega_{\rm t}}\delta\vartheta)$. Then there exists a constant $c > 0$ independent on $\mathbf{q}$ and $\mathbf{r}$ such that*

$$\begin{aligned} &\|\boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{V}^p(\Omega)} + \|\varpi_{\mathbf{q},\mathbf{r}} - \varpi_{\mathbf{r}}\|_{L^p(\Omega)/\mathbb{R}} + \|\eta_{\mathbf{q},\mathbf{r}} - \eta_{\mathbf{r}}\|_{W_0^{1,q}(\Omega)} \\ &\qquad \leq c(\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)})(\|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{r}}\|_{L^2(\Omega)}) \end{aligned} \tag{4.4}$$

$$\|\boldsymbol{v}_{\mathbf{q},\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q},\mathbf{r}}\|_{L^2(\Omega)} \leq c(\|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{r}}\|_{L^2(\Omega)}). \tag{4.5}$$

Proof. First, let us observe from the definitions of $(\boldsymbol{v}_{\mathbf{q},\mathbf{r}}, \varpi_{\mathbf{q},\mathbf{r}}, \eta_{\mathbf{q},\mathbf{r}})$ and $(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}})$ that

$$A_{(\boldsymbol{u}_{\mathbf{q}},P_{\mathbf{q}},\theta_{\mathbf{q}})}(\boldsymbol{v}_{\mathbf{q},\mathbf{r}}, \varpi_{\mathbf{q},\mathbf{r}}, \eta_{\mathbf{q},\mathbf{r}}) = (\chi_{\omega_{\mathrm{f}}}\delta\boldsymbol{\mu}, \chi_{\omega_{\mathrm{t}}}\delta\vartheta)$$
$$A_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})}(\boldsymbol{v}_{\mathbf{r}}, \varpi_{\mathbf{r}}, \eta_{\mathbf{r}}) = (\chi_{\omega_{\mathrm{f}}}\delta\boldsymbol{\mu}, \chi_{\omega_{\mathrm{t}}}\delta\vartheta).$$

Let $\boldsymbol{v} = \boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}}$, $\varpi = \varpi_{\mathbf{q},\mathbf{r}} - \varpi_{\mathbf{r}}$ and $\eta = \eta_{\mathbf{q},\mathbf{r}} - \eta_{\mathbf{r}}$. A simple calculation reveals that

$$A_{(\boldsymbol{u}_{\mathbf{q}},P_{\mathbf{q}},\theta_{\mathbf{q}})}(\boldsymbol{v}, \varpi, \eta) = (\boldsymbol{f}, h) \tag{4.6}$$

where $\boldsymbol{f} = -[(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}})\cdot\nabla]\boldsymbol{v}_{\mathbf{r}} - [\boldsymbol{v}_{\mathbf{r}}\cdot\nabla](\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}})$ and $h = -[(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}})\cdot\nabla]\eta_{\mathbf{r}} - [\boldsymbol{v}_{\mathbf{r}}\cdot\nabla](\theta_{\mathbf{q}} - \bar{\theta})$.

Given $\boldsymbol{\psi} \in \boldsymbol{W}_0^{1,p'}(\Omega)$, upon applying $\boldsymbol{V}^p(\Omega) \subset \boldsymbol{L}^{2p/(p-2)}(\Omega)$, we get

$$\begin{aligned}|\langle \boldsymbol{f}, \boldsymbol{\psi}\rangle| &\le c\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^{2p/(2-p)}(\Omega)}\|\nabla\boldsymbol{\psi}\|_{\boldsymbol{L}^{p'}(\Omega)^2}\|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)}\\ &\le c\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{V}^p(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p'}(\Omega)}\|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)}.\end{aligned}$$

On the other hand, given $\phi \in W_0^{1,q'}(\Omega)$ and using the assumption $q \le p$, we have $\boldsymbol{V}^p(\Omega) \subset \boldsymbol{V}^q(\Omega) \subset \boldsymbol{L}^{2q/(2-q)}(\Omega)$ and thus

$$\begin{aligned}|\langle h, \phi\rangle| &\le (\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^{2q/(2-q)}(\Omega)}\|\eta_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)}\|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^{2q/(2-q)}(\Omega)})\|\nabla\phi\|_{\boldsymbol{L}^{q'}(\Omega)}\\ &\le (\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{V}^p(\Omega)}\|\eta_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)}\|\theta_{\mathbf{q}} - \bar{\theta}\|_{W_0^{1,q}(\Omega)})\|\phi\|_{W_0^{1,q'}(\Omega)}.\end{aligned}$$

Therefore, it follows from the above calculations that

$$\begin{aligned}&\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|h\|_{W^{-1,q}(\Omega)}\\ &\qquad \le c(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{V}^p(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{W_0^{1,q}(\Omega)})(\|\boldsymbol{v}_{\mathbf{r}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{r}}\|_{L^2(\Omega)}).\end{aligned} \tag{4.7}$$

The inequality (4.4) now follows from (4.2), (4.3), (4.6) and (4.7). Finally, (4.5) is a consequence of (4.4), the triangle inequality and the continuity of $\boldsymbol{V}^p(\Omega) \times W_0^{1,q}(\Omega) \subset \boldsymbol{L}^2(\Omega) \times L^2(\Omega)$. □

Next, we estimate the L^2-distance of the solutions with controls in $\mathcal{U}$ to the optimal state.

Lemma 4.3. *There exist $\varepsilon > 0$ and $c_\varepsilon > 0$ such that if $(\boldsymbol{v}_{\mathbf{q}}, \varpi_{\mathbf{q}}, \eta_{\mathbf{q}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\mathrm{t}}}\bar{\vartheta})(\chi_{\omega_{\mathrm{f}}}\boldsymbol{\mu}, \chi_{\omega_{\mathrm{t}}}\vartheta)$ and $(\boldsymbol{v}_{\bar{\mathbf{q}}}, \varpi_{\bar{\mathbf{q}}}, \eta_{\bar{\mathbf{q}}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\mathrm{t}}}\bar{\vartheta})(\chi_{\omega_{\mathrm{f}}}\bar{\boldsymbol{\mu}}, \chi_{\omega_{\mathrm{t}}}\bar{\vartheta})$, then*

$$\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)} \le c_\varepsilon(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)}) \tag{4.8}$$

for all $\mathbf{q} = (\boldsymbol{\mu}, \vartheta) \in \mathcal{U}$ such that $\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)} < \varepsilon$.

Proof. Let $\boldsymbol{u} = \boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}} - \boldsymbol{v}_{\mathbf{q}} + \boldsymbol{v}_{\bar{\mathbf{q}}}$, $P = P - P_{\mathbf{q}} - \varpi_{\mathbf{q}} + \varpi_{\bar{\mathbf{q}}}$ and $\theta = \theta_{\mathbf{q}} - \bar{\theta} - \eta_{\mathbf{q}} + \eta_{\bar{\mathbf{q}}}$. Then one can see that

$$A_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})}(\boldsymbol{u}, P, \theta) = -([(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}})\cdot\nabla](\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}), [(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}})\cdot\nabla](\theta_{\mathbf{q}} - \bar{\theta})).$$

Applying a similar argument as in the previous lemma, one has

$$\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|\theta\|_{W_0^{1,q}(\Omega)} \le c\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{V}^p(\Omega)}(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}). \tag{4.9}$$

Therefore, from the Lipschitz estimate (4.2), the triangle inequality and inequality (4.9)

$$\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}$$

$$\begin{aligned}
&\le \|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta\|_{L^2(\Omega)} + \|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)} \\
&\le c(\|\boldsymbol{u}\|_{\boldsymbol{V}^p(\Omega)} + \|\theta\|_{W_0^{1,q}(\Omega)}) + \|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)} \\
&\le c\varepsilon(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}) + \|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)},
\end{aligned}$$

whenever $\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)} < \varepsilon$. Choosing $\varepsilon > 0$ small enough so that $c\varepsilon < 1$, we get the estimate (4.8) with the constant $c_\varepsilon = (1 - c\varepsilon)^{-1}$. □

In the following lemma, we shall estimate the residual norms for the adjoint states.

Lemma 4.4. *There exists $c > 0$ such that for all $\mathbf{q} \in \mathcal{U}$ we have*

$$\|\boldsymbol{\varphi}_{\mathbf{q}} - \bar{\boldsymbol{\varphi}}\|_{\boldsymbol{V}^{p'}(\Omega)} + \|\zeta_{\mathbf{q}} - \bar{\zeta}\|_{W_0^{1,q'}(\Omega)} \le c(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}) \tag{4.10}$$

where $(\boldsymbol{\varphi}_{\mathbf{q}}, \pi_{\mathbf{q}}, \zeta_{\mathbf{q}})$ is the solution of (3.3).

Proof. Let $(\boldsymbol{\varphi}, \pi, \zeta) = (\boldsymbol{\varphi}_{\mathbf{q}} - \bar{\boldsymbol{\varphi}}, \pi_{\mathbf{q}} - \bar{\pi}, \zeta_{\mathbf{q}} - \bar{\zeta}_{\mathbf{q}})$. A straightforward calculation shows that $A^\star_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})}(\boldsymbol{\varphi}, \pi, \zeta) = (\boldsymbol{f}, h)$ where

$$\begin{aligned}
h &= \theta_{\mathbf{q}} - \bar{\theta} + (\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}) \cdot \nabla\zeta_{\mathbf{q}} \\
\boldsymbol{f} &= \boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}} + (\theta_{\mathbf{q}} - \bar{\theta})\nabla\zeta_{\mathbf{q}} + [(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}) \cdot \nabla]\boldsymbol{\varphi}_{\mathbf{q}} - (\nabla\boldsymbol{u}_{\mathbf{q}} - \nabla\bar{\boldsymbol{u}})^T\boldsymbol{\varphi}_{\mathbf{q}}.
\end{aligned}$$

The adjoint states $(\boldsymbol{\varphi}_{\mathbf{q}}, \pi_{\mathbf{q}}, \zeta_{\mathbf{q}})$ are uniformly bounded for $\mathbf{q} \in \mathcal{U}$, *i.e.*

$$\|\boldsymbol{\varphi}_{\mathbf{q}}\|_{\boldsymbol{V}^{p'}(\Omega)} + \|\pi_{\mathbf{q}}\|_{L^{p'}(\Omega)/\mathbb{R}} + \|\zeta_{\mathbf{q}}\|_{W_0^{1,q'}(\Omega)} \le c(\|\boldsymbol{u}_{\mathbf{q}} - \boldsymbol{u}_d\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \theta_d\|_{L^2(\Omega)}) \le c$$

thanks to (4.3), the triangle inequality and the boundedness of $\mathcal{U}$.

Let us estimate the norms of $\boldsymbol{f}$ and h in $\boldsymbol{W}^{-1,p'}(\Omega)$ and $W^{-1,q'}(\Omega)$, respectively. Concerning the last two terms in $\boldsymbol{f}$, take an arbitrary $\boldsymbol{\psi} \in \boldsymbol{W}_0^{1,p}(\Omega)$ and proceed as follows:

$$\begin{aligned}
&|\langle[(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}) \cdot \nabla]\boldsymbol{\varphi}_{\mathbf{q}} - (\nabla\boldsymbol{u}_{\mathbf{q}} - \nabla\bar{\boldsymbol{u}})^T\boldsymbol{\varphi}_{\mathbf{q}}, \boldsymbol{\psi}\rangle| \\
&\qquad \le |\langle[(\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}) \cdot \nabla]\boldsymbol{\varphi}_{\mathbf{q}}, \boldsymbol{\psi}\rangle| + |\langle(\boldsymbol{\psi} \cdot \nabla)\boldsymbol{\varphi}_{\mathbf{q}}, \boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\rangle| \\
&\qquad \le c\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{\varphi}_{\mathbf{q}}\|_{\boldsymbol{V}^{p'}(\Omega)}\|\nabla\boldsymbol{\psi}\|_{\boldsymbol{L}^{2p/(2-p)}(\Omega)^2} \\
&\qquad \le c\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{\varphi}_{\mathbf{q}}\|_{\boldsymbol{V}^{p'}(\Omega)}\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)}.
\end{aligned}$$

According to the continuity of $W_0^{1,q'}(\Omega) \subset W_0^{1,p'}(\Omega)$ and $\boldsymbol{L}^2(\Omega) \subset \boldsymbol{W}^{-1,p'}(\Omega)$, we have

$$\begin{aligned}
&|\langle\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}} + (\theta_{\mathbf{q}} - \bar{\theta})\nabla\zeta_{\mathbf{q}}, \boldsymbol{\psi}\rangle| \\
&\qquad \le c(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{W}^{-1,p'}(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}\|\zeta_{\mathbf{q}}\|_{W_0^{1,p'}(\Omega)})\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)} \\
&\qquad \le c(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)}\|\zeta_{\mathbf{q}}\|_{W_0^{1,q'}(\Omega)})\|\boldsymbol{\psi}\|_{\boldsymbol{W}_0^{1,p}(\Omega)}.
\end{aligned}$$

Combining the above estimates and using the boundedness of the adjoint states yield

$$\|\boldsymbol{f}\|_{\boldsymbol{W}^{-1,p'}(\Omega)} \le c(\|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)} + \|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)}).$$

A similar process shows that

$$\|h\|_{W^{-1,q}(\Omega)} = \|\theta_{\mathbf{q}} - \bar{\theta} + (\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}) \cdot \nabla\zeta_{\mathbf{q}}\|_{W^{-1,q}(\Omega)}$$

$$\leq c(\|\theta_{\mathbf{q}} - \bar{\theta}\|_{L^2(\Omega)} + \|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)}).$$

The desired estimate (4.10) now follows from the fact that $A^\star_{(\bar{\boldsymbol{u}},\bar{P},\bar{\theta})}$ is an isomorphism from $\boldsymbol{V}^{p'}(\Omega) \times (L^{p'}(\Omega)/\mathbb{R}) \times W^{1,q'}_0(\Omega)$ onto $\boldsymbol{W}^{-1,p'}(\Omega) \times W^{-1,q'}(\Omega)$, *cf.* Lemma 2.15. □

We shall now estimate the second derivatives of the smooth part of the cost functional with directions that are deviations of the optimal control.

Lemma 4.5. *For every $\rho > 0$ there exists $\varepsilon_\rho > 0$ such that*

$$|[j_d''(\boldsymbol{\mu},\vartheta) - j_d''(\bar{\boldsymbol{\mu}},\bar{\vartheta})](\boldsymbol{\mu}-\bar{\boldsymbol{\mu}},\vartheta-\bar{\vartheta})^2| \leq \rho(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|^2_{L^2(\Omega)}) \tag{4.11}$$

for every $\mathbf{q} = (\boldsymbol{\mu},\vartheta) \in \mathcal{U}$ satisfying $\|\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)} < \varepsilon_\rho$.

Proof. Let us denote the deviation in controls by $\mathbf{r} = (\delta\boldsymbol{\mu},\delta\vartheta) = (\boldsymbol{\mu}-\bar{\boldsymbol{\mu}},\vartheta-\bar{\vartheta})$. Further, write $(\boldsymbol{v}_{\mathbf{q},\mathbf{r}},\varpi_{\mathbf{q},\mathbf{r}},\eta_{\mathbf{q},\mathbf{r}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\rm f}}\boldsymbol{\mu}, h_d + \chi_{\omega_{\rm t}}\vartheta)(\chi_{\omega_{\rm f}}\delta\boldsymbol{\mu},\chi_{\omega_{\rm t}}\delta\vartheta)$ and $(\boldsymbol{v}_{\mathbf{r}},\varpi_{\mathbf{r}},\eta_{\mathbf{r}}) = \mathcal{S}'(\boldsymbol{f}_d + \chi_{\omega_{\rm f}}\bar{\boldsymbol{\mu}}, h_d + \chi_{\omega_{\rm t}}\bar{\vartheta})(\chi_{\omega_{\rm f}}\delta\boldsymbol{\mu},\chi_{\omega_{\rm t}}\delta\vartheta)$. According to the representation of j_d'' in (3.2), we have

$$\begin{aligned}
&|[j_d''(\boldsymbol{\mu},\vartheta) - j_d''(\bar{\boldsymbol{\mu}},\bar{\vartheta})](\delta\boldsymbol{\mu},\delta\vartheta)^2| \\
&\quad\leq \int_\Omega |\boldsymbol{v}_{\mathbf{q},\mathbf{r}} + \boldsymbol{v}_{\mathbf{r}}||\boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}}|\,\mathrm{d}x + \int_\Omega |\eta_{\mathbf{q},\mathbf{r}} + \eta_{\mathbf{r}}||\eta_{\mathbf{q},\mathbf{r}} - \eta_{\mathbf{r}}|\,\mathrm{d}x \\
&\quad+ 2\int_\Omega \{|[(\boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}})\cdot\nabla]\boldsymbol{\varphi}_{\mathbf{q}}\cdot\boldsymbol{v}_{\mathbf{q},\mathbf{r}}| + |(\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)(\boldsymbol{\varphi}_{\mathbf{q}} - \bar{\boldsymbol{\varphi}})\cdot\boldsymbol{v}_{\mathbf{q},\mathbf{r}}| + |(\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)\bar{\boldsymbol{\varphi}}\cdot(\boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}})|\}\,\mathrm{d}x \\
&\quad+ 2\int_\Omega \{|[(\boldsymbol{v}_{\mathbf{q},\mathbf{r}} - \boldsymbol{v}_{\mathbf{r}})\cdot\nabla]\zeta_{\mathbf{q}}\eta_{\mathbf{q},\mathbf{r}}| + |(\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)(\zeta_{\mathbf{q}} - \bar{\zeta})\eta_{\mathbf{q},\mathbf{r}}| + |(\boldsymbol{v}_{\mathbf{r}}\cdot\nabla)\bar{\zeta}(\eta_{\mathbf{q},\mathbf{r}} - \eta_{\mathbf{r}})|\}\,\mathrm{d}x.
\end{aligned}$$

The integral terms can be estimated from above by following the same strategy as in the proof of Lemma 4.2. Invoking Lemmas 4.2–4.4, the embedding $\boldsymbol{V}^p(\Omega)\times W^{1,q}_0(\Omega) \subset \boldsymbol{L}^2(\Omega)\times L^2(\Omega)$ and $(\boldsymbol{v}_{\mathbf{r}},\eta_{\mathbf{r}}) = (\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}},\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}})$ it can be verified that

$$\begin{aligned}
&|[j_d''(\boldsymbol{\mu},\vartheta) - j_d''(\bar{\boldsymbol{\mu}},\bar{\vartheta})](\delta\boldsymbol{\mu},\delta\vartheta)^2| \\
&\quad\leq c(\|\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta-\bar{\vartheta}\|_{W^{-1,q}(\Omega)})(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|^2_{L^2(\Omega)}).
\end{aligned}$$

Given $\rho > 0$, we take $\varepsilon_\rho > 0$ small enough so that $c\varepsilon_\rho < \rho$. Therefore, we have (4.11) whenever $\|\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta-\bar{\vartheta}\|_{W^{-1,q}(\Omega)} < \varepsilon_\rho$. □

With the same procedure as in the above lemma, one can deduce the following estimate.

Lemma 4.6. *There exists $c_0 > 0$ such that for all $\mathbf{q} = (\boldsymbol{\mu},\vartheta) \in \mathcal{U}$ we have*

$$\begin{aligned}
&|j_d''(\boldsymbol{\mu},\vartheta)(\boldsymbol{\mu}-\bar{\boldsymbol{\mu}},\vartheta-\bar{\vartheta})^2| \\
&\quad\leq c_0(\|\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta-\bar{\vartheta}\|_{W^{-1,q}(\Omega)})(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)}).
\end{aligned}$$

We are now in position to state and prove the sufficient optimality conditions for the optimal control problem (1.1)–(1.3).

Theorem 4.7. *Let $4/3 < q \leq p < 2$. Suppose that $(\bar{\boldsymbol{\mu}},\bar{\vartheta})$ satisfies the local first-order necessary optimality conditions as stated in Theorem 3.2. Assume that there exist $\tau > 0$ and $\kappa > 0$ such that*

$$j_d''(\bar{\boldsymbol{\mu}},\bar{\vartheta})(\boldsymbol{\mu},\vartheta)^2 \geq \kappa(\|\boldsymbol{v}_{\mathbf{q}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}}\|^2_{L^2(\Omega)}), \qquad \forall \mathbf{q} = (\boldsymbol{\mu},\vartheta) \in \mathcal{C}_\tau(\bar{\boldsymbol{\mu}},\bar{\vartheta}). \tag{4.12}$$

Then there exist constants $\varepsilon = \varepsilon_{\tau,\kappa} > 0$ *and* $\gamma = \gamma_{\tau,\kappa} > 0$ *such that*

$$j(\bar{\boldsymbol{\mu}}, \bar{\vartheta}) + \gamma(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|^2_{L^2(\Omega)}) \le j(\boldsymbol{\mu}, \vartheta) \tag{4.13}$$

for all $\mathbf{q} = (\boldsymbol{\mu}, \vartheta) \in \mathcal{U}$ *satisfying* $\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\boldsymbol{W}^{-1,p}(\Omega)} + \|\vartheta - \bar{\vartheta}\|_{W^{-1,q}(\Omega)} < \varepsilon$.

Proof. As in [8], let us distinguish two cases. First, suppose that $(\delta\boldsymbol{\mu}, \delta\vartheta) := (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}, \vartheta - \bar{\vartheta}) \in \mathcal{C}_\tau(\bar{\boldsymbol{u}}, \bar{\vartheta})$. Note that from (3.1), $-\frac{1}{\alpha}\bar{\boldsymbol{\varphi}} \in \partial j_{\mathrm{f}}(\bar{\boldsymbol{\mu}}) \cap \boldsymbol{C}_0(\omega_{\mathrm{f}})$ and $-\frac{1}{\beta}\bar{\zeta} \in \partial j_{\mathrm{t}}(\bar{\vartheta}) \cap C_0(\omega_{\mathrm{t}})$, we have

$$j'((\bar{\boldsymbol{\mu}}, \bar{\vartheta}); (\delta\boldsymbol{\mu}, \delta\vartheta)) = \int_{\omega_{\mathrm{f}}} \bar{\boldsymbol{\varphi}}\, \mathrm{d}(\delta\boldsymbol{\mu}) + \alpha j_{\mathrm{f}}'(\bar{\boldsymbol{\mu}}; \delta\mu) + \int_{\omega_{\mathrm{t}}} \bar{\zeta}\, \mathrm{d}(\delta\vartheta) + \beta j_{\mathrm{t}}'(\bar{\vartheta}; \delta\vartheta) \ge 0.$$

Applying a Taylor expansion, Lemma 4.3, Lemma 4.5 with $0 < \rho < \kappa$ and (4.12), we have for some $\sigma \in [0, 1]$ that

$$\begin{aligned}
&j(\boldsymbol{\mu}, \vartheta) - j(\bar{\boldsymbol{\mu}}, \bar{\vartheta}) \\
&\quad \ge \frac{1}{2} j_d''(\bar{\boldsymbol{\mu}}, \bar{\vartheta})(\delta\boldsymbol{\mu}, \delta\vartheta)^2 + \frac{1}{2}[j_d''(\bar{\boldsymbol{\mu}} + \sigma(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}), \bar{\vartheta} + \sigma(\vartheta - \bar{\vartheta})) - j_d''(\bar{\boldsymbol{\mu}}, \bar{\vartheta})](\delta\boldsymbol{\mu}, \delta\vartheta)^2 \\
&\quad \ge \frac{1}{2}(\kappa - \rho)(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|^2_{L^2(\Omega)}) \\
&\quad \ge \frac{1}{2c_\varepsilon}(\kappa - \rho)(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|^2_{L^2(\Omega)}).
\end{aligned}$$

On the other hand, if $(\delta\boldsymbol{\mu}, \delta\vartheta) \notin \mathcal{C}_\tau(\bar{\boldsymbol{u}}, \bar{\vartheta})$, then

$$j'(\bar{\boldsymbol{\mu}}, \bar{\vartheta})(\delta\boldsymbol{\mu}, \delta\vartheta) > \tau(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)}).$$

By choosing $\varepsilon \in (0, \varepsilon_\rho)$ small enough, one has $\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|_{L^2(\Omega)} \le 1$ thanks to (4.2) and (4.3). A Taylor expansion once more, together with Lemma 4.6, leads to

$$\begin{aligned}
j(\boldsymbol{\mu}, \vartheta) - j(\bar{\boldsymbol{\mu}}, \bar{\vartheta}) &\ge j'(\bar{\boldsymbol{\mu}}, \bar{\vartheta})(\delta\boldsymbol{\mu}, \delta\vartheta) + \frac{1}{2} j_d''(\bar{\boldsymbol{\mu}} + \sigma(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}), \bar{\vartheta} + \sigma(\vartheta - \bar{\vartheta}))(\delta\boldsymbol{\mu}, \delta\vartheta)^2 \\
&\ge \frac{1}{2}(\tau - c_0\varepsilon)(\|\boldsymbol{v}_{\mathbf{q}} - \boldsymbol{v}_{\bar{\mathbf{q}}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\eta_{\mathbf{q}} - \eta_{\bar{\mathbf{q}}}\|^2_{L^2(\Omega)}) \\
&\ge \frac{1}{2c_\varepsilon}(\tau - c_0\varepsilon)(\|\boldsymbol{u}_{\mathbf{q}} - \bar{\boldsymbol{u}}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\theta_{\mathbf{q}} - \bar{\theta}\|^2_{L^2(\Omega)}).
\end{aligned}$$

Reducing $\varepsilon > 0$ further if necessary in such a way that $c_0\varepsilon < \tau$, we obtain (4.13) with $\gamma = \min\{\kappa - \rho, \tau - c_0\varepsilon\}/(2c_\varepsilon) > 0$. This completes the proof of the theorem. ☐

5. Numerical approximations

In this section, we consider a finite element approximation for the optimal control problem (1.1)–(1.3) based on a semi-smooth Newton method developed in [5, 6, 20, 24]. Let $\{\mathcal{T}_h\}_{h>0}$ be a family of quasi-regular triangulations of Ω parametrized by their mesh sizes h, that is, the length of the largest triangular edge. We shall only present the case of the triangular mini-finite element space [3], and furthermore, for simplicity of exposition we shall consider the case $\omega_{\mathrm{f}} = \omega_{\mathrm{t}} = \Omega$, $\boldsymbol{f}_d = 0$ and $h_d = 0$. The discussions below can be adapted to the case of triangular Taylor–Hood elements as well as for quadrilateral basis functions, see [13] for instance, as long as the discrete inf-sup condition holds.

5.1. Finite element approximation

Let $\{x_i\}_{i=1}^{N_h}$ be the set of all nodes together with the barycenters of the triangles in the mesh. Denote by I_{ph} the set of all nodes in the triangulation and I_h to be the set of nodes together with the barycenters. Associated with these nodes, we consider the nodal Lagrange basis functions $\{e_i\}_{i\in I_h}$, consisting of continuous piecewise linear polynomials for vertex nodes and bubble functions for the barycenters, such that $e_i(x_j) = \delta_{ij}$ for all $i, j \in I_h$. Similarly, we denote the linear Lagrange basis elements $\{\lambda_i\}_{i\in I_{ph}}$ so that $\lambda_i(x_j) = \delta_{ij}$ for every $i, j \in I_{ph}$.

Consider the finite element spaces

$$V_h = \left\{ u_h \in C(\bar\Omega) : u_h = \sum_{i\in I_h} u_{hi} e_i,\ u_{hi} \in \mathbb{R} \right\}$$
$$Q_h = \left\{ p_h \in C(\bar\Omega) : p_h = \sum_{i\in I_{ph}} p_{hi} \lambda_i,\ p_{hi} \in \mathbb{R} \right\}.$$

Let $\boldsymbol{V}_h = V_h \times V_h$ and denote $\boldsymbol{e}_{i,j} = (e_i, e_j)$ for $i, j \in I_h$. For the discretization of the control space, we consider the following space of linear combinations of Dirac measures concentrated on the nodes and barycenters

$$D_h = \left\{ \mu_h \in M(\Omega) : \mu_h = \sum_{i\in I_h} \mu_{hi} \delta_{x_i},\ \mu_{hi} \in \mathbb{R} \right\}$$

and let $\boldsymbol{D}_h = D_h \times D_h$. Note that $\boldsymbol{D}_h$ can be identified with the dual of $\boldsymbol{V}_h$. For $u_h, v_h \in V_h$ and $\delta_{x_i} \in D_h$ we define $\langle u_h \delta_{x_i}, v_h\rangle := u_h(x_i)v_h(x_i)$ and similarly for the vector-valued case.

To approximate the optimal solutions, we follow the strategy of optimize-then-discretize, that is, we discretize the optimality system for the continuous problem. For the state equation, consider the nonlinear operator

$$T : \boldsymbol{W}_0^{1,p}(\Omega) \times (L^p(\Omega)/\mathbb{R}) \times W_0^{1,q}(\Omega) \to \boldsymbol{W}^{-1,p}(\Omega) \times L^p(\Omega) \times W^{-1,q}(\Omega)$$
$$T(\boldsymbol{u}, P, \theta) = \begin{pmatrix} -\nu\Delta\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla P - \eta\boldsymbol{g} \\ \operatorname{div}\boldsymbol{u} \\ -\kappa\Delta\eta + (\boldsymbol{u}\cdot\nabla)\theta \end{pmatrix}.$$

On the the other hand, for the adjoint equation let us define the linear operator

$$L : \boldsymbol{W}_0^{1,p'}(\Omega) \times (L^{p'}(\Omega)/\mathbb{R}) \times W_0^{1,q'}(\Omega) \to \boldsymbol{W}^{-1,p'}(\Omega) \times L^{p'}(\Omega) \times W^{-1,q'}(\Omega)$$
$$L(\boldsymbol{\varphi}, \pi, \zeta) = \begin{pmatrix} -\nu\Delta\boldsymbol{\varphi} + (\nabla\boldsymbol{u})^T\boldsymbol{\varphi} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{\varphi} + \nabla\pi - \theta\nabla\zeta \\ \operatorname{div}\boldsymbol{\varphi} \\ -\kappa\Delta\zeta - (\boldsymbol{u}\cdot\nabla)\zeta - \boldsymbol{g}\cdot\boldsymbol{\varphi} \end{pmatrix}.$$

From Section 3, the optimality system for (1.1)–(1.3) can be equivalently written as

$$\begin{cases} T(\bar{\boldsymbol{u}}, \bar P, \bar\theta) = (\bar{\boldsymbol{\mu}}, 0, \bar\vartheta) \\ L(\bar{\boldsymbol{\varphi}}, \bar\pi, \bar\zeta) = (\bar{\boldsymbol{u}} - \boldsymbol{u}_d, 0, \bar\theta - \theta_d) \\ \langle \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\varphi}} - \boldsymbol{\varphi}\rangle_{\boldsymbol{M}(\Omega)\times\boldsymbol{C}_0(\Omega)} \le 0 \qquad \forall\, \|\boldsymbol{\varphi}\|_{\boldsymbol{C}_0(\Omega)} \le \alpha \\ \langle \bar\vartheta, \bar\zeta - \zeta\rangle_{M(\Omega)\times C_0(\Omega)} \le 0 \qquad \forall\, \|\zeta\|_{C_0(\Omega)} \le \beta. \end{cases}$$

Discretizing the above continuous optimality system, reformulating the corresponding variational inequalities in terms of the max and min functions and adding a Moreau–Yosida regularization, we obtain the discrete system (see [10] for the details)

$$\begin{cases} T_h(\bar{\boldsymbol{u}}_h, \bar{P}_h, \bar{\theta}_h) = (\bar{\boldsymbol{\mu}}_h, 0, \bar{\vartheta}_h) \\ L_h(\bar{\boldsymbol{\varphi}}_h, \bar{\pi}_h, \bar{\zeta}_h) = (\bar{\boldsymbol{u}}_h - \boldsymbol{u}_{dh}, 0, \bar{\theta}_h - \theta_{dh}) \\ \bar{\boldsymbol{\mu}}_h + \max(0, -\bar{\boldsymbol{\mu}}_h + \gamma(\bar{\boldsymbol{\varphi}}_h - \alpha)) + \min(0, -\bar{\boldsymbol{\mu}}_h + \gamma(\bar{\boldsymbol{\varphi}}_h + \alpha)) = 0 \\ \bar{\vartheta}_h + \max(0, -\bar{\vartheta}_h + \gamma(\bar{\zeta}_h - \beta)) + \min(0, -\bar{\vartheta}_h + \gamma(\bar{\zeta}_h + \beta)) = 0. \end{cases} \tag{5.1}$$

Here, T_h and L_h are the finite element discretization of T and L, respectively, and will be discussed below. Also, $\boldsymbol{u}_{dh} \in \boldsymbol{V}_h$ and $\theta_{dh} \in V_h$ are the Lagrange approximations of the desired states $\boldsymbol{u}_d$ and θ_d.

The goal of the parameter $\gamma > 0$ is to provide a starting point for the semi-smooth Newton method when applied to the above discretized optimality system. First, we initialize $\gamma > 0$ and set $\bar{\boldsymbol{\mu}}_h = 0$ and $\bar{\vartheta}_h = 0$ inside of the min and max functions. The resulting system can be expressed as follows:

$$\begin{cases} T_h(\boldsymbol{u}_h, p_h, \theta_h) = (\boldsymbol{\mu}_h, 0, \vartheta_h) \\ L_h(\boldsymbol{\varphi}_h, \pi_h, \zeta_h) = (\boldsymbol{u}_h - \boldsymbol{u}_{dh}, 0, \theta_h - \theta_{dh}) \\ \boldsymbol{\mu}_h + \gamma\chi_{\mathcal{A}(\boldsymbol{\varphi}_h)}\boldsymbol{\varphi}_h - \gamma\alpha(\chi_{\mathcal{A}_+(\boldsymbol{\varphi}_h)} - \chi_{\mathcal{A}_-(\boldsymbol{\varphi}_h)}) = 0 \\ \vartheta_h + \gamma\chi_{\mathcal{A}(\zeta_h)}\zeta_h - \gamma\beta(\chi_{\mathcal{A}_+(\zeta_h)} - \chi_{\mathcal{A}_-(\zeta_h)}) = 0, \end{cases}$$

with the active sets of indices

$$\begin{aligned} &\mathcal{A}_+(\boldsymbol{\varphi}_h) := \{i \in I_h : \boldsymbol{\varphi}_h(x_i) > \alpha\}, && \mathcal{A}_+(\zeta_h) := \{i \in I_h : \zeta_h(x_i) > \beta\}, \\ &\mathcal{A}_-(\boldsymbol{\varphi}_h) := \{i \in I_h : \boldsymbol{\varphi}_h(x_i) < -\alpha\}, && \mathcal{A}_-(\zeta_h) := \{i \in I_h : \zeta_h(x_i) < -\beta\}, \\ &\mathcal{A}(\boldsymbol{\varphi}_h) := \mathcal{A}_+(\boldsymbol{\varphi}_h) \cup \mathcal{A}_-(\boldsymbol{\varphi}_h), && \mathcal{A}(\zeta_h) := \mathcal{A}_+(\zeta_h) \cup \mathcal{A}_-(\zeta_h). \end{aligned}$$

This system is solved iteratively by evaluating the indicator functions at previous values of the adjoint variables. More precisely, starting initially with $\boldsymbol{\varphi}_h^0 = 0$ and $\zeta_h^0 = 0$, the update at the kth step is given by

$$\begin{cases} T_h(\boldsymbol{u}_h^k, p_h^k, \theta_h^k) = (\boldsymbol{\mu}_h^k, 0, \vartheta_h^k) \\ L_h(\boldsymbol{\varphi}_h^k, \pi_h^k, \zeta_h^k) = (\boldsymbol{u}_h^k - \boldsymbol{u}_{dh}, 0, \theta_h^k - \theta_{dh}) \\ \boldsymbol{\mu}_h^k + \gamma\chi_{\mathcal{A}(\boldsymbol{\varphi}_h^{k-1})}\boldsymbol{\varphi}_h^k - \gamma\alpha(\chi_{\mathcal{A}_+(\boldsymbol{\varphi}_h^{k-1})} - \chi_{\mathcal{A}_-(\boldsymbol{\varphi}_h^{k-1})}) = 0 \\ \vartheta_h^k + \gamma\chi_{\mathcal{A}(\zeta_h^{k-1})}\zeta_h^k - \gamma\beta(\chi_{\mathcal{A}_+(\zeta_h^{k-1})} - \chi_{\mathcal{A}_-(\zeta_h^{k-1})}) = 0 \end{cases} \tag{5.2}$$

until the stopping criteria $\mathcal{A}(\boldsymbol{\varphi}_h^k) = \mathcal{A}(\boldsymbol{\varphi}_h^{k-1})$ and $\mathcal{A}(\zeta_h^k) = \mathcal{A}(\zeta_h^{k-1})$ are satisfied. From the last two equations, we see that the variables $\boldsymbol{\mu}_h^k$ and ϑ_h^k can be eliminated from the system. We then repeat calculating the solution of (5.2), but now with γ replaced by $\sigma\gamma$ for some scaling factor $\sigma > 1$, with the solution of the previous problem as the initial iterate. This process is terminated once the stopping criteria is satisfied or a declared maximum parameter γ has been reached.

Let $\boldsymbol{\varphi}_h^{(0)}$ and $\zeta_h^{(0)}$ denote the solution of the above procedure and γ^* be the final value of γ. For the semi-smooth Newton method, we take the initial point

$$\begin{aligned} \boldsymbol{\mu}_h^{(0)} &= -\gamma^*[\max(0, \boldsymbol{\varphi}_h^{(0)} - \alpha) + \min(0, \boldsymbol{\varphi}_h^{(0)} + \alpha)] \\ \vartheta_h^{(0)} &= -\gamma^*[\max(0, \zeta_h^{(0)} - \beta) + \min(0, \zeta_h^{(0)} + \beta)]. \end{aligned}$$

This choice of initialization is based on (5.1). Given discrete controls $\boldsymbol{\mu}_h^{(j-1)}$ and $\vartheta_h^{(j-1)}$, we then solve the coupled state and adjoint system

$$\begin{cases} T_h(\boldsymbol{u}_h^{(j)}, p_h^{(j)}, \theta_h^{(j)}) = (\boldsymbol{\mu}_h^{(j-1)}, 0, \vartheta_h^{(j-1)}) \\ L_h(\boldsymbol{\varphi}_h^{(j)}, \pi_h^{(j)}, \zeta_h^{(j)}) = (\boldsymbol{u}_h^{(j)} - \boldsymbol{u}_{dh}, 0, \theta_h^{(j)} - \theta_{dh}) \end{cases} \tag{5.3}$$

and update the control according to

$$\begin{aligned} \boldsymbol{\mu}_h^{(j)} &= -\chi_{\mathcal{A}(\boldsymbol{\varphi}_h^{(j-1)} - \boldsymbol{\mu}_h^{(j-1)})}(\boldsymbol{\varphi}_h^{(j-1)} - \boldsymbol{\mu}_h^{(j-1)}) + \alpha(\chi_{\mathcal{A}_+(\boldsymbol{\varphi}_h^{(j-1)} - \boldsymbol{\mu}_h^{(j-1)})} - \chi_{\mathcal{A}_-(\boldsymbol{\varphi}_h^{(j-1)} - \boldsymbol{\mu}_h^{(j-1)})}) \\ \vartheta_h^{(j)} &= -\chi_{\mathcal{A}(\zeta_h^{(j-1)} - \vartheta_h^{(j-1)})}(\zeta_h^{(j-1)} - \vartheta_h^{(j-1)}) + \beta(\chi_{\mathcal{A}_+(\zeta_h^{(j-1)} - \vartheta_h^{(j-1)})} - \chi_{\mathcal{A}_-(\zeta_h^{(j-1)} - \vartheta_h^{(j-1)})}). \end{aligned}$$

Again, if there are no more changes on the active sets then this subroutine is terminated, that is, when $\mathcal{A}(\boldsymbol{\varphi}_h^{(j)} - \boldsymbol{\mu}_h^{(j)}) = \mathcal{A}(\boldsymbol{\varphi}_h^{(j-1)} - \boldsymbol{\mu}_h^{(j-1)})$ and $\mathcal{A}(\zeta_h^{(j)} - \vartheta_h^{(j)})) = \mathcal{A}(\zeta_h^{(j-1)} - \vartheta_h^{(j-1)})$.

The coupled discretized nonlinear state and adjoint equations (5.2) are solved by Newton's method. Moreover, for stabilization purposes, we add an artificial compressibility penalty parameter $0 < \varepsilon \ll 1$. Let

$$\mathcal{X}_h := \boldsymbol{D}_h \times Q_h \times D_h \times \boldsymbol{D}_h \times Q_h \times D_h.$$

The nonlinear finite-dimensional system corresponding to (5.2) is given by

$$F_{h\varepsilon}(X_h^k) := F_{h\varepsilon}(\boldsymbol{u}_h^k, p_h^k, \theta_h^k, \boldsymbol{\varphi}_h^k, \pi_h^k, \zeta_h^k) = 0 \tag{5.4}$$

where $F_{h\varepsilon} : \mathcal{X}_h \to \mathcal{X}_h'$ is defined by

$$\begin{aligned} &\langle F_{h\varepsilon}(X_h^k), Y_h\rangle_{\mathcal{X}_h' \times \mathcal{X}_h} \\ &:= \int_\Omega \{\nu \nabla \boldsymbol{u}_h^k \cdot \nabla \boldsymbol{v}_h + (\boldsymbol{u}_h^k \cdot \nabla)\boldsymbol{u}_h^k \cdot \boldsymbol{v}_h - p_h^k \operatorname{div} \boldsymbol{v}_h - \theta_h^k \boldsymbol{g} \cdot \boldsymbol{v}_h\} \, \mathrm{d}x \\ &- \int_\Omega \varpi_h \operatorname{div} \boldsymbol{u}_h^k \, \mathrm{d}x + \int_\Omega \varepsilon p_h^k \varpi_h \, \mathrm{d}x + \int_\Omega \{\kappa \nabla \theta_h^k \cdot \nabla \eta_h + (\boldsymbol{u}_h^k \cdot \nabla)\theta_h^k \eta_h\} \, \mathrm{d}x \\ &+ \int_\Omega \{\nu \nabla \boldsymbol{\varphi}_h^k \cdot \nabla \boldsymbol{\psi}_h - (\boldsymbol{u}_h^k \cdot \nabla)\boldsymbol{\varphi}_h^k \cdot \boldsymbol{\psi}_h + (\nabla \boldsymbol{u}_h^k)^T \boldsymbol{\varphi}_h^k \cdot \boldsymbol{\psi}_h - \pi_h^k \operatorname{div} \boldsymbol{\psi}_h\} \, \mathrm{d}x \\ &- \int_\Omega \{\theta_h^k \nabla \zeta_h^k \cdot \boldsymbol{\psi}_h + (\boldsymbol{u}_h^k - \boldsymbol{u}_{dh}) \cdot \boldsymbol{\psi}_h\} \, \mathrm{d}x - \int_\Omega \rho_h \operatorname{div} \boldsymbol{\varphi}_h^k \, \mathrm{d}x + \int_\Omega \varepsilon \pi_h^k \rho_h \, \mathrm{d}x \\ &+ \int_\Omega \{\kappa \nabla \zeta_h^k \cdot \nabla \phi_h - (\boldsymbol{u}_h^k \cdot \nabla)\zeta_h^k \phi_h - \boldsymbol{g} \cdot \boldsymbol{\varphi}_h^k \phi_h - (\theta_h^k - \theta_{dh})\phi_h\} \, \mathrm{d}x \\ &+ \sum_{i \in \mathcal{A}(\boldsymbol{\varphi}_h^k)} \gamma \langle \boldsymbol{\varphi}_h^k \boldsymbol{\delta}_{x_i}, \boldsymbol{v}_h\rangle - \sum_{i \in \mathcal{A}_+(\boldsymbol{\varphi}_h^k)} \alpha\gamma \langle \boldsymbol{\delta}_{x_i}, \boldsymbol{v}_h\rangle + \sum_{i \in \mathcal{A}_-(\boldsymbol{\varphi}_h^k)} \alpha\gamma \langle \boldsymbol{\delta}_{x_i}, \boldsymbol{v}_h\rangle \\ &+ \sum_{i \in \mathcal{A}(\zeta_h^k)} \gamma \langle \zeta_h \delta_{x_i}, \eta_h\rangle - \sum_{i \in \mathcal{A}_+(\zeta_h^k)} \beta\gamma \langle \delta_{x_i}, \eta_h\rangle + \sum_{i \in \mathcal{A}_-(\zeta_h^k)} \beta\gamma \langle \delta_{x_i}, \eta_h\rangle \end{aligned}$$

for every $Y_h := (\boldsymbol{v}_h, \varpi_h, \eta_h, \boldsymbol{\psi}_h, \rho_h, \phi_h) \in \mathcal{X}_h$.

Approximating a solution to (5.4) by Newton's method requires the Jacobian of $F_{h\varepsilon}$. However, instead of directly taking the Jacobian of $F_{h\varepsilon}$, we take an inexact approach by calculating the derivative of the functional associated with the continuous system, applying the properties of the trilinear forms induced by the convection

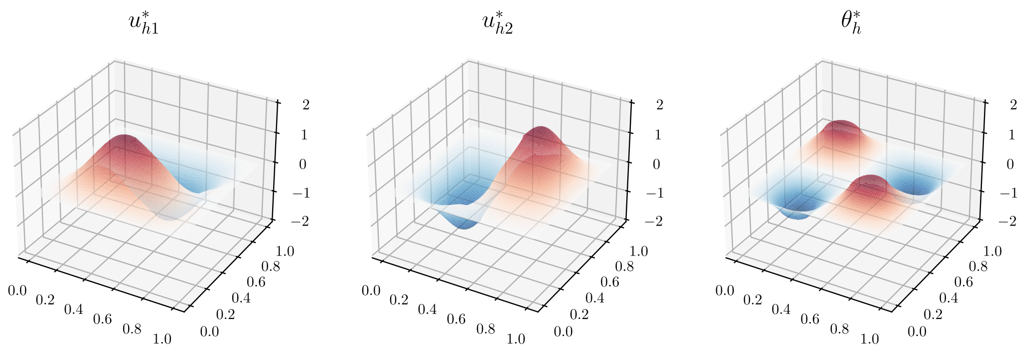

FIGURE 1. Components of the numerical optimal velocity $\boldsymbol{u}_h^* = (u_{h1}^*, u_{h2}^*)$ and temperature θ_h^*.

terms and consider the finite-dimensional approximation. The linear operator $G_{h\varepsilon} : \mathcal{X}_h \to \mathcal{X}_h'$ obtained from this procedure is given by

$$\begin{aligned}
&\langle G_{h\varepsilon}(X_h^k)\bar{X}_h, Y_h\rangle_{\mathcal{X}_h'\times\mathcal{X}_h} \\
&\quad := \int_\Omega \{\nu\nabla\bar{\boldsymbol{u}}_h \cdot \nabla\boldsymbol{v}_h + (\nabla\boldsymbol{u}_h^k)^T\bar{\boldsymbol{u}}_h \cdot \boldsymbol{v}_h + (\boldsymbol{u}_h^k \cdot \nabla)\bar{\boldsymbol{u}}_h \cdot \boldsymbol{v}_h - \bar{p}_h \operatorname{div}\boldsymbol{v}_h - \bar{\theta}_h\boldsymbol{g}\cdot\boldsymbol{v}_h\}\,\mathrm{d}x \\
&\quad - \int_\Omega \varpi_h \operatorname{div}\bar{\boldsymbol{u}}_h\,\mathrm{d}x + \int_\Omega \varepsilon\bar{p}_h\varpi_h\,\mathrm{d}x + \int_\Omega \{\kappa\nabla\bar{\theta}_h\cdot\nabla\eta_h + (\bar{\boldsymbol{u}}_h\cdot\nabla)\theta_h^k\eta_h + (\boldsymbol{u}_h^k\cdot\nabla)\bar{\theta}_h\eta_h\}\,\mathrm{d}x \\
&\quad + \int_\Omega \{\nu\nabla\bar{\boldsymbol{\varphi}}_h\cdot\nabla\boldsymbol{\psi}_h - (\boldsymbol{u}_h^k\cdot\nabla)\bar{\boldsymbol{\varphi}}_h\cdot\boldsymbol{\psi}_h - [\nabla\boldsymbol{\varphi}_h^k + (\nabla\boldsymbol{\varphi}_h^k)^T]\bar{\boldsymbol{u}}_h\cdot\boldsymbol{\psi}_h + (\nabla\boldsymbol{u}_h^k)^T\bar{\boldsymbol{\varphi}}_h\cdot\boldsymbol{\psi}_h\}\,\mathrm{d}x \\
&\quad - \int_\Omega \{\bar{\pi}_h \operatorname{div}\boldsymbol{\psi}_h + \bar{\theta}_h\nabla\zeta_h^k\cdot\boldsymbol{\psi}_h + \theta_h^k\nabla\bar{\zeta}_h\cdot\boldsymbol{\psi}_h + \bar{\boldsymbol{u}}_h\cdot\boldsymbol{\psi}_h\}\,\mathrm{d}x - \int_\Omega \rho_h \operatorname{div}\bar{\boldsymbol{\varphi}}_h\,\mathrm{d}x \\
&\quad + \int_\Omega \varepsilon\bar{\pi}_h\rho_h\,\mathrm{d}x + \int_\Omega \{\kappa\nabla\bar{\zeta}_h\cdot\nabla\phi_h - (\bar{\boldsymbol{u}}_h\cdot\nabla)\zeta_h^k\phi_h - (\boldsymbol{u}_h^k\cdot\nabla)\bar{\zeta}_h\phi_h - \boldsymbol{g}\cdot\bar{\boldsymbol{\varphi}}_h\phi_h - \bar{\theta}_h\phi_h\}\,\mathrm{d}x \\
&\quad + \sum_{i\in\mathcal{A}(\bar{\boldsymbol{\varphi}}_h)} \gamma\langle\bar{\boldsymbol{\varphi}}_h\boldsymbol{\delta}_{x_i}, \boldsymbol{v}_h\rangle + \sum_{i\in\mathcal{A}(\bar{\zeta}_h)} \gamma\langle\bar{\zeta}_h\delta_{x_i}, \eta_h\rangle
\end{aligned}$$

where $\bar{X}_h := (\bar{\boldsymbol{u}}_h, \bar{p}_h, \bar{\theta}_h, \bar{\boldsymbol{\varphi}}_h, \bar{\pi}_h, \bar{\zeta}_h) \in \mathcal{X}_h$. Given $X_h^{k,j} \in \mathcal{X}_h$ we compute the solution $\bar{X}_h^{k,j} \in \mathcal{X}_h$ of the linear system

$$G_{h\varepsilon}(X_h^{k,j})\bar{X}_h^{k,j} = -F_{h\varepsilon}(X_h^{k,j}) \tag{5.5}$$

and update according to $X_h^{k,j+1} = X_h^{k,j} - \bar{X}_h^{k,j}$. This subroutine is terminated once we have $\|X_h^{k,j}\|_{\mathcal{X}_h} < \tau$ for some prescribed tolerance $0 < \tau \ll 1$. The above *inexact* Newton iterative scheme is initialized by $X_h^{k,0} = 0$. This procedure is also adapted in the approximation of solutions to the discretized nonlinear primal-dual system (5.3).

5.2. Numerical example

For the computational domain, we take the unit square $\Omega = (0,1)\times(0,1)$. We consider a uniform triangulation of Ω consisting of 10201 nodes and 20000 triangles, which corresponds to a mesh size $h = \sqrt{2}/100$. The parameters in the cost functional are chosen as $\alpha = \beta = 10^{-3}$ and $\nu = \kappa = 1$. We set $\gamma = \sigma = 10$ in the continuation strategy. For the desired states we consider $u_{d1}(x,y) = (1-\cos(2\pi x))\sin(2\pi y)$, $u_{d2}(x,y) = (1-\cos(2\pi y))\sin(2\pi x)$ and $\theta_d(x,y) = -\sin(2\pi x)\sin(2\pi y)$. Note that $\boldsymbol{u}_d = (u_{d1}, u_{d2})$ is divergence-free in Ω.

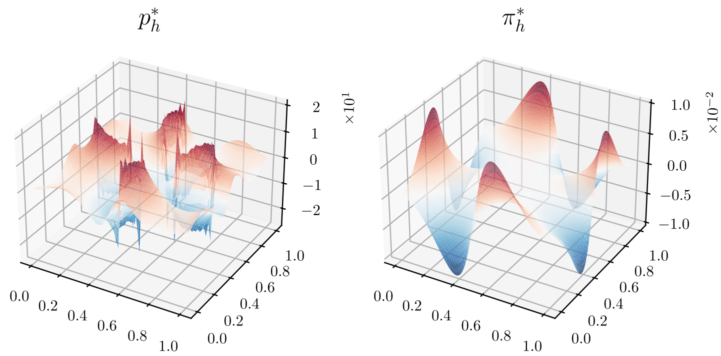

FIGURE 2. Numerical optimal pressure p_h^* and adjoint pressure π_h^*.

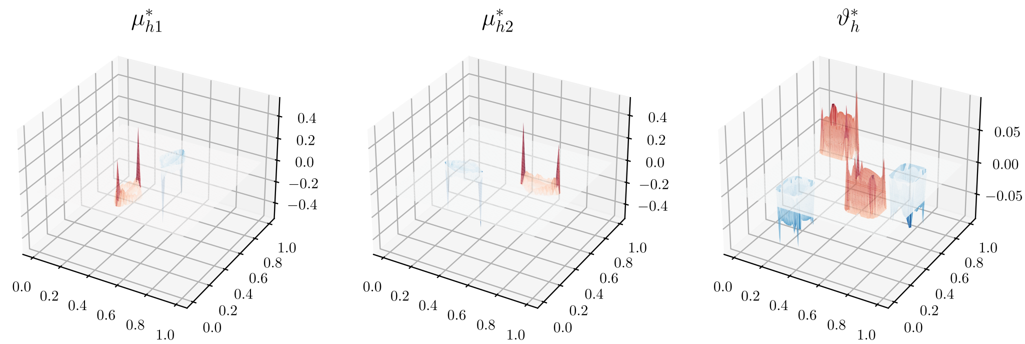

FIGURE 3. Linear interpolation of numerical velocity controls $\boldsymbol{\mu}_h^* = (\mu_{h1}^*, \mu_{h2}^*)$ and thermal optimal control ϑ_h^*.

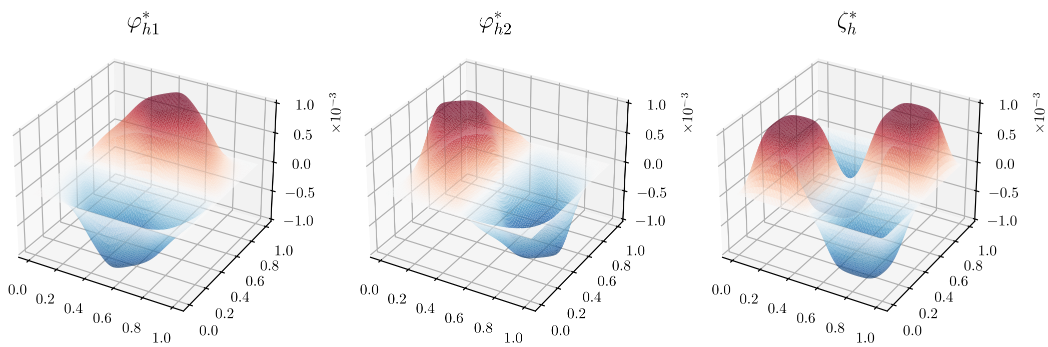

FIGURE 4. The components of the numerical optimal adjoint velocity $\boldsymbol{\varphi}_h^* = (\varphi_{h1}^*, \varphi_{h2}^*)$ and adjoint temperature ζ_h^*.

The algorithm described in the previous subsection was implemented in Python 3.9.7 (Python Software Foundation, https://www.python.org/) on a 2.3 GHz Intel Core i5 with 8 GB RAM. The source codes and iteration histories can be downloaded at https://github.com/grperalta/boussinesqmeasure. Regarding the linear system (5.5), we take a penalty parameter $\varepsilon = 10^{-11}$. Each of the matrices appearing on both sides of this system were assembled at every Newton iteration using algorithms analogous to those provided in Chapters 7 and 8 of [11] with Gaussian quadrature of order 6. In this case, the total number of degrees of freedom corresponding to the primal and dual variables is $\dim \mathcal{X}_h = 201608$. The solutions of the linearized primal-dual systems (5.5) were obtained by utilizing the sparse solver splu in the python package SciPy with the UMFPACK option and terminated once $\|X_h^{k,j}\| \leq 10^{-10}$. The discrete system (5.3) is solved successively with $\gamma = 10^k$ for $0 \leq k \leq 7$.

The optimal states corresponding to the penalty parameter $\gamma^* = 10^7$ are given in Figures 1 and 2. Likewise, the optimal controls are presented in Figure 3, where a linear interpolation was utilized for better visualization. Both the velocity and thermal controls have sparse supports, and the symmetric or anti-symmetric properties of the desired states about the center of Ω are reflected as well in the optimal controls.

In Figure 4, the profiles of the numerical optimal adjoint states are shown, and we have $\|\zeta_h^*\|_\infty \approx 10^{-3}$ and $\|\boldsymbol{\varphi}_h^*\|_\infty = \max\{\|\varphi_{h1}^*\|_\infty, \|\varphi_{h2}^*\|_\infty\} \approx 10^{-3}$. As predicted from Theorem 3.3, the supports of the positive and negative parts of the discrete controls $\mu_{h1}^{*\pm}$, $\mu_{h2}^{*\pm}$ and $\vartheta_h^{*\pm}$ are located on the nodes where $\varphi_{h1}^* \approx \mp 10^{-3}$, $\varphi_{h2}^* \approx \mp 10^{-3}$ and $\zeta_h^* \approx \mp 10^{-3}$, respectively.

Acknowledgements. This work was supported in part by the One U.P. Faculty Grant 2019-101374. The author is grateful to the referee for the valuable comments and suggestions that led to the enhancement of the manuscript.

References

[1] F. Abergel and R. Temam, On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dyn.* **1** (1990) 303–325.
[2] R.A. Adams, Sobolev Spaces. Academic Press, New York (1975).
[3] D.N. Arnold, F. Brezzi, and M. Fortin, A stable finite element for the Stokes equations. *Calcolo* **21** (1984) 337-344.
[4] J.L. Boldrini, E. Fernandéz-Cara and M. A. Rojas-Medar, An optimal control problem for a generalized Boussinesq model: the time dependent case. *Rev. Mat. Complete.* **20** (2007) 339–366.
[5] E. Casas, C. Clason and K. Kunisch, Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **50** (2012) 1735-1752.
[6] E. Casas, C. Clason and K. Kunisch, Parabolic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **51** (2013) 28-63.
[7] E. Casas and K. Kunisch, Optimal control of semilinear elliptic equations in measure spaces. *SIAM J. Control Optim.* **52** (2014) 339-364.
[8] E. Casas and K. Kunisch, Optimal control of the two-dimensional stationary Navier–Stokes equations with measure valued controls. *SIAM J. Control Optim.* **57** (2019) 1328-1354.
[9] E. Casas and K. Kunisch, Parabolic control problems in space-time measure spaces. *ESAIM: COCV* **22** (2016) 355-370.
[10] C. Clason and K. Kunisch, A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM: COCV* **17** (2011) 243-266.
[11] A. Ern and J.-L. Guermond, Theory and Practice of Finite Elements. Springer (2004).
[12] G. Galdi, An Introduction to the Mathematical Theory of the Navier–Stokes Equations. Springer, New York (2011).
[13] V. Girault and P. A. Raviart, Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms. Springer-Verlag, Berlin (1986).
[14] M. Hinze and U. Matthes, Optimal and model predictive control of the Boussinesq approximation, in Control of Coupled Partial Differential Equations, edited by K. Kunisch, J. Sprekels, G. Leugering and F. Tröltzsch. Birkhäuser Basel (2007).
[15] K. Ito and S.S. Ravindran, Optimal control of thermally convected fluid flows. *SIAM J. Sci. Comput.* **19** (1998) 1847–1869.
[16] H. Kim, Existence and regularity of very weak solutions of the stationary Navier–Stokes equations. *Arch. Ration. Mech. Anal.* **193** (2009) 117-152.
[17] H. Kim, The existence and uniqueness of very weak solutions of the stationary Boussinesq system. *Nonlinear Anal. Real World Appl.* **75** (2012) 317-330.
[18] K. Kunisch, K. Pieper and B. Vexler, Measure valued directional sparsity for parabolic optimal control problems. SIAM J. Control Optim. **52** (2014) 3078–3108.
[19] K. Kunisch, P. Trautmann and B. Vexler, Optimal control of the undamped linear wave equation with measure valued controls. *SIAM J. Control Optim.* **54** (2016) 1212-1244.
[20] B. Kummer, Newton's method based on generalized derivatives for nonsmooth functions: Convergence analysis, edited by Oettli W., Pallaschke D. Advances in Optimization. *Lecture Notes in Economics and Mathematical Systems*, vol 382. Springer, Berlin, Heidelberg (1992) 171-194.
[21] H.-C. Lee and B. C. Shin, Piecewise optimal distributed controls for 2D Boussinesq equations. *Math. Methods Appl. Sci.* **23** (2000) 227-254.
[22] H.-C. Lee and O.Yu. Imanuvilov, Analysis of optimal control problems for the 2-D stationary Boussinesq equations. *J. Math. Anal. Appl.* **242** (2000) 191-211.
[23] J. M. Mihaljan, A rigorous exposition of the Boussinesq approximations applicable to a thin layer of fluid. *Astrophys. J.* **136** (1962) 1126-1133.
[24] L. Qi and J. Sun, A nonsmooth version of Newton's method. *Math. Program.* **58** (1993) 353-367.
[25] W. Rudin, Real and Complex Analysis. McGraw-Hill Book Co., London (1970).
[26] P.A. Tapia, *L^p-theory for the Boussinesq system*, Ph.D. thesis, University of Chile (2015).

[27] F. Tröltzsch and D. Wachsmuth, Second-order sufficient optimality conditions for the optimal control of Navier–Stokes equations. *ESAIM: COCV* **12** (2006) 353-367.
[28] E. Zeidler, Nonlinear Functional Analysis and its Applications I, Springer-Verlag, New York (1986).