# Optimal Control Problems for the Two Dimensional Rayleigh-Bénard Type Convection by a Gradient Method 

Hyung-Chun Lee<br>Department of Mathematics, Ajou University<br>Suwon 443-749, Korea

Received February 8, 2007
Revised October 28, 2008


#### Abstract

In this aricle, the author considers mathematical formulation and numerical solutions of distributed and Neumann boundary optimal control problems associated with the stationary Bénard problem. The solution of the optimal control problem is obtained by controlling of the source term of the equations and/or Neumann boundary conditions. Then the author considers the approximation, by finite element methods, of the optimality system and derive optimal error estimates. The convergence of a simple gradient method is proved and some numerical results are given.


Key words: flow control, temperature control, Boussinesq equations, optimization

## 1. Introduction

Rayleigh-Bénard convection is a convective flow when a stagnant flow layer is heated from below such that its low side is hotter than its upper. This problem involves buoyancy driven flow in a cavity. In this article, a stabilization of two dimensional Rayleigh-Bénard type convection is considered. The boundary condition for the problem involved two horizontal walls which differ in temperature leading to a thermal gradient across the solution domain. The flow is heated from a part $\left(\Gamma_{1}\right)$ of the bottom boundary. The boundary conditions for the problem are described in Fig. 1. In the present investigation, we consider methods suppressing the natural convection by adjusting heat flux distribution at a part $\left(\Gamma_{N}\right)$ of the bottom or adjusting heat source on the flow domain $\Omega$.

In past years, considerable progress has been made in mathematical analyses and computations of optimal control problems for viscous flows; see [2, 4, 5, $9,11,12,13,14,18,19,20,21,22,26,27]$ and references therein. Optimal control problems for the thermally coupled incompressible Navier-Stokes equation by Neumann and Diriclet boundary heat controls were considered in [18, 22]. Also, optimal control problems for the time dependent optimal control problems for the Bénard problem and related problems were considered in [5, 9, 26] and references therein. Exact controllability of the Boussinesq problem and related problems were considered in [14].

For the Navier-Stokes cases, many computational methods such as SQP, Newton-type second order methods, trust-region methods, and the conjugate gradient method were proposed in $[3,10,15,23,24]$ and references therein. Even

[^0]

Fig. 1. Computational domain.
though the gradient type methods are slow in convergence, it is still convenient to use because of its simplicity and stability, especially for the complex system. Of course, the fast second-order schemes have some drawbacks such as stability and complexity.

The problem we consider is a Bénard problem whose system is governed by the Boussinesq equations. We now write the 2-D nondimensional Boussinesq equations as follows:

$$
\left\{\begin{array}{l}
-\operatorname{Pr} \Delta \boldsymbol{u}+(\boldsymbol{u} \cdot \nabla) \boldsymbol{u}=-\nabla p+R \operatorname{Pr} T \boldsymbol{j} \quad \text { in } \Omega,  \tag{1.1}\\
\nabla \cdot \boldsymbol{u}=0 \quad \text { in } \Omega \\
-\Delta T+(\boldsymbol{u} \cdot \nabla) T=Q \quad \text { in } \Omega
\end{array}\right.
$$

with boundary conditions

$$
\left\{\begin{array}{l}
\boldsymbol{u}=\mathbf{0} \quad \text { on } \partial \Omega, \quad T=g_{D} \quad \text { on } \Gamma_{1}, \quad T=0 \quad \text { on } \Gamma_{\mathrm{top}},  \tag{1.2}\\
\frac{\partial T}{\partial \boldsymbol{n}}=g_{N} \quad \text { on } \Gamma_{N}, \quad \frac{\partial T}{\partial \boldsymbol{n}}=0 \quad \text { on } \Gamma_{\text {left }} \cup \Gamma_{\text {right }}
\end{array}\right.
$$

where the computational domain $\Omega=(0,1) \times(0,1) \subset \mathbb{R}^{2}$, with Lipschitz boundary $\partial \Omega=\Gamma_{\text {top }} \cup \Gamma_{\text {bottom }} \cup \Gamma_{\text {left }} \cup \Gamma_{\text {right }}$. Here $\Gamma_{\text {bottom }}=\Gamma_{1} \cup \Gamma_{N} ;$ see Fig. 1. In (1.1)-(1.2), $\boldsymbol{u}, p$ and $T$ denote the velocity, pressure and temperature fields, respectively, $g_{D}$ a given function, and $Q$ and $g_{N}$ controls. The vector $\boldsymbol{j}$ is a unit vector in the direction of gravitational acceleration. The dimensionless variables are defined by the following relations, where the superscript asterisk denotes dimensional quantities:

$$
\begin{align*}
& x=\frac{x^{*}}{d_{x}}, \quad y=\frac{y^{*}}{d_{y}}, \quad \boldsymbol{u}=\frac{d_{y} \boldsymbol{u}^{*}}{\kappa}, \quad R=\alpha g \frac{T_{H}^{*}-T_{C}^{*}}{\kappa \nu} d_{y}^{3}, \\
& \operatorname{Pr}=\frac{\nu}{\kappa}, \quad T=\frac{T^{*}-T_{C}^{*}}{T_{H}^{*}-T_{C}^{*}}, \quad P^{\prime}:=\frac{d_{y}^{2} p^{*}}{\rho \kappa^{2}}, \tag{1.3}
\end{align*}
$$

where $\kappa>0$ the thermal conductivity parameter, $\rho$ the density parameter, $d_{x}$ the half width of the cavity, $d_{y}$ the half depth of the domain, $T_{H}^{*}$ the nominal bottom temperature, $T_{C}^{*}$ the temperature at the top boundary, $R$ the Rayleigh number, $\operatorname{Pr}$ the Prandtl number, $\alpha$ the thermal expansion coefficient, $\nu$ the kinematic viscosity and $g$ the gravitational constant. In this paper we consider, for the simplicity, the case of constant $\kappa$ and constant $\nu$. The vector $\boldsymbol{n}$ denotes the outward unit normal to $\Omega$. The modified pressure $p$ is given by

$$
\begin{equation*}
p=P^{\prime}-\left(T_{C}^{*}-T_{\mathrm{ref}}^{*}\right) \frac{d_{y}^{3}}{\kappa^{2}} \alpha g y, \tag{1.4}
\end{equation*}
$$

where $T_{\text {ref }}^{*}$ is the reference temperature given by

$$
\begin{equation*}
T_{\mathrm{ref}}^{*}=\frac{1}{2}\left(T_{H}^{*}+T_{C}^{*}\right) \tag{1.5}
\end{equation*}
$$

Next, we introduce the functionals

$$
\begin{equation*}
\mathcal{J}_{1}\left(\boldsymbol{u}, T, p, Q, g_{N}\right)=\frac{1}{2 \alpha} \int_{\Omega}|\nabla \times \boldsymbol{u}|^{2} d \boldsymbol{x}+\frac{\beta}{2} \int_{\Omega}|Q|^{2} d \boldsymbol{x}+\frac{\gamma}{2} \int_{\Gamma_{N}}\left|g_{N}\right|^{2} d s \tag{1.6}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathcal{J}_{2}\left(\boldsymbol{u}, T, p, Q, g_{N}\right)=\frac{1}{2 \alpha} \int_{\Omega}\left|\boldsymbol{u}-\boldsymbol{U}_{d}\right|^{2} d \boldsymbol{x}+\frac{\beta}{2} \int_{\Omega}|Q|^{2} d \boldsymbol{x}+\frac{\gamma}{2} \int_{\Gamma_{N}}\left|g_{N}\right|^{2} d s \tag{1.7}
\end{equation*}
$$

The positive penalty parameters $\alpha, \beta$ and $\gamma$ can be used to change the relative importance of the three terms appearing in the definitions of the functionals. They can also be used as regularization parameters. The optimal control problems we consider are to seek state variables $(\boldsymbol{u}, T, p)$, and controls $Q$ and $g_{N}$ such that the functional (1.6) or (1.7) is minimized subject to (1.1)-(1.2) where $\boldsymbol{U}_{d}$ is some desired velocity field. The functional (1.6) measures the vorticity of the flow. The control of vorticity has significant applications in science and engineering such as control of turbulence and control of crystal growth process. The first term in the functional (1.7) measures the $L^{2}$-distance between the candidate flow and the desired flow. Thus, the physical objective of this minimization problem is to match a desired flow field (in the $L^{2}$-sense) by adjusting the distributed control $Q$ or the boundary temperature flux $g_{N}$. The real goal of optimization is to minimize the first term appearing in the definition (1.6)-(1.7). The second and third terms in the cost functionals (1.6)-(1.7) are added to limit the cost of controls.

The plan of the paper is as follows. In $\S 1.1$ of this section we introduce the notation that will be used throughout the paper. In $\S 2$, we introduce a weak formulation of mathematical problem and derive the optimality system of the optimal control problem. Then, we introduce finite element approximations and derive optimal error estimates. In $\S 3$, a gradient method for the solution of the discrete optimal
control problem is introduced and the convergence of the gradient method is proved. Finally in $\S 4$, the results of some computational experiments are presented. The computational results show that our method is very efficient and feasible.

### 1.1. Notation

We introduce some function spaces and their norms, along with some related notations used in subsequent sections; for details see [1].

Let $\Omega$ be a bounded domain of $\mathbb{R}^{2}$ with a Lipschitz boundary $\partial \Omega$. Let $L^{2}(\Omega)$ be the space of real-valued square integrable functions defined on $\Omega$, and let $\|\cdot\|_{L^{2}(\Omega)}$ be the norm in this space. We define the Sobolev space $H^{m}(\Omega)$ for the nonnegative integer $m$ by

$$
H^{m}(\Omega)=\left\{u \in L^{2}(\Omega) \mid D^{a} u \in L^{2}(\Omega), \text { for } 0 \leq|a| \leq m\right\},
$$

where $D^{a}$ is the weak (or distributional) partial derivative, $a$ is a multi-index. The norm $\|\cdot\|_{H^{m}(\Omega)}$ associated with $H^{m}(\Omega)$ is given by

$$
\|u\|_{H^{m}(\Omega)}^{2}=\sum_{|a| \leq m}\left\|D^{a} u\right\|_{L^{2}(\Omega)}^{2} .
$$

Note that $H^{0}(\Omega)=L^{2}(\Omega)$. For the vector-valued functions, we define the Sobolev space $\boldsymbol{H}^{m}(\Omega)$ (in all cases, boldface indicates vector-valued) by

$$
\boldsymbol{H}^{m}(\Omega)=\left\{\boldsymbol{u}=\left(u_{1}, u_{2}\right) \mid u_{i} \in H^{m}(\Omega), \text { for } i=1,2\right\}
$$

and its associated norm $\|\cdot\|_{\boldsymbol{H}^{m}(\Omega)}$ is given by

$$
\|\boldsymbol{u}\|_{\boldsymbol{H}^{m}(\Omega)}^{2}=\sum_{i=1}^{2}\left\|u_{i}\right\|_{H^{m}(\Omega)}^{2} .
$$

We also define particular subspaces:

$$
L_{0}^{2}(\Omega)=\left\{f \in L^{2}(\Omega) \mid \int_{\Omega} f d \boldsymbol{x}=0\right\}, \quad \boldsymbol{H}_{0}^{1}(\Omega)=\left\{\boldsymbol{u} \in \boldsymbol{H}^{1}(\Omega) \mid \boldsymbol{u}=\mathbf{0} \text { on } \Gamma\right\},
$$

and

$$
H_{D}^{1}(\Omega)=\left\{S \in H^{1}(\Omega) \mid S=0 \text { on } \Gamma_{D}\right\}
$$

where $\Gamma_{D}=\Gamma_{1} \cup \Gamma_{\mathrm{top}}$. We also define the solenoidal space

$$
\boldsymbol{V}=\left\{\boldsymbol{u} \in \boldsymbol{H}_{0}^{1}(\Omega) \mid \nabla \cdot \boldsymbol{u}=0\right\} .
$$

## 2. Preliminaries

### 2.1. A weak formulation of the equations

We introduce the following bilinear and trilinear forms, for all $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w} \in$ $\boldsymbol{H}^{1}(\Omega), T, S \in H^{1}(\Omega)$ and $q \in L_{0}^{2}(\Omega)$,

$$
\begin{array}{cl}
a_{0}(\boldsymbol{u}, \boldsymbol{v})=\int_{\Omega} \nabla \boldsymbol{u}: \nabla \boldsymbol{v} d \boldsymbol{x}, & a_{1}(T, S)=\int_{\Omega} \nabla T \cdot \nabla S d \boldsymbol{x}, \\
c_{0}(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v})=\int_{\Omega}(\boldsymbol{u} \cdot \nabla) \boldsymbol{w} \cdot \boldsymbol{v} d \boldsymbol{x}, \quad c_{1}(\boldsymbol{u}, T, S)=\int_{\Omega}(\boldsymbol{u} \cdot \nabla) T S d \boldsymbol{x},
\end{array}
$$

and

$$
b(\boldsymbol{v}, q)=-\int_{\Omega} q \nabla \cdot \boldsymbol{v} d \boldsymbol{x}, \quad d(T, \boldsymbol{v})=R \operatorname{Pr} \int_{\Omega} T \boldsymbol{j} \cdot \boldsymbol{v} d \boldsymbol{x}
$$

We first note that the bilinear forms $a_{0}(\cdot, \cdot)$ and $a_{1}(\cdot, \cdot)$ are clearly continuous, i.e.,

$$
\begin{align*}
&\left|a_{0}(\boldsymbol{u}, \boldsymbol{v})\right| \leq C\|\boldsymbol{u}\|_{\boldsymbol{H}^{1}(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{H}^{1}(\Omega)},  \tag{2.1}\\
&\left|a_{1}(T, S)\right| \leq C\|T\|_{H^{1}(\Omega)}\|S\|_{H^{1}(\Omega)}, \tag{2.2}
\end{align*}
$$

and

$$
\begin{equation*}
|b(\boldsymbol{v}, q)| \leq C\|\boldsymbol{v}\|_{\boldsymbol{H}^{1}(\Omega)}\|q\|_{L^{2}(\Omega)} . \tag{2.3}
\end{equation*}
$$

We have the coercivity relations associated with $a_{0}(\cdot, \cdot)$ and $a_{1}(\cdot, \cdot)$ :

$$
\begin{equation*}
a_{0}(\boldsymbol{u}, \boldsymbol{u})=\|\nabla \boldsymbol{u}\|_{L^{2}(\Omega)}^{2} \geq C_{1}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1}(\Omega)}^{2} \quad \forall \boldsymbol{u} \in \boldsymbol{H}_{0}^{1}(\Omega) \tag{2.4}
\end{equation*}
$$

and

$$
\begin{equation*}
a_{1}(T, T)=\|\nabla T\|_{L^{2}(\Omega)}^{2} \geq C_{2}\|T\|_{H^{1}(\Omega)}^{2} \quad \forall T \in H_{D}^{1}(\Omega), \tag{2.5}
\end{equation*}
$$

which are direct consequences of Poincaré inequality.
Lemma 2.1. For every $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v} \in \boldsymbol{H}^{1}(\Omega)$ and every $T, S \in H^{1}(\Omega)$ there are constants $C_{1}$ and $C_{2}$ such that

$$
\begin{gather*}
\left|c_{0}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v})\right| \leq C_{1}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1}(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{H}^{1}(\Omega)}\|\boldsymbol{v}\|_{\boldsymbol{H}^{1}(\Omega)},  \tag{2.6}\\
c_{0}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v})=0 \quad \text { if } \boldsymbol{u} \in \boldsymbol{V},  \tag{2.7}\\
\left|c_{1}(\boldsymbol{u}, T, S)\right| \leq C_{2}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1}(\Omega)}\|T\|_{H^{1}(\Omega)}\|S\|_{H^{1}(\Omega)} \quad \forall \boldsymbol{u} \in \boldsymbol{V}, \tag{2.8}
\end{gather*}
$$

and

$$
\begin{equation*}
c_{1}(\boldsymbol{u}, T, T)=0 \quad \text { if } \boldsymbol{u} \in \boldsymbol{V} \tag{2.9}
\end{equation*}
$$

Proof. These follow from the Cauchy-Schwarz inequality, Hölder's inequality, and various embedding results, in particular the continuous embeddings of $\boldsymbol{H}^{1}$ into $\boldsymbol{L}^{4}$ and $\boldsymbol{L}^{2}$ and $H^{1}$ into $L^{4}$ and $L^{2}$, respectively.

The weak form of the constraint equations (1.1)-(1.2) is then given as follows: seek $\boldsymbol{u} \in \boldsymbol{H}_{0}^{1}(\Omega), p \in L_{0}^{2}(\Omega)$ and $T \in H^{1}(\Omega)$ such that

$$
\begin{align*}
\operatorname{Pr} a_{0}(\boldsymbol{u}, \boldsymbol{v})+c_{0}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})+b(\boldsymbol{v}, p)=d(T, \boldsymbol{v}) & \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega),  \tag{2.10}\\
b(\boldsymbol{u}, q)=0 \quad \forall q \in L_{0}^{2}(\Omega), &  \tag{2.11}\\
a_{1}(T, S)+c_{1}(\boldsymbol{u}, T, S)=\langle Q, S\rangle-\left\langle g_{N}, S\right\rangle_{\Gamma_{N}} & \forall S \in H_{D}^{1}(\Omega), \tag{2.12}
\end{align*}
$$

and

$$
\begin{equation*}
T=g_{D} \quad \text { on } \Gamma_{1}, \quad T=0 \quad \text { on } \Gamma_{\mathrm{top}}, \tag{2.13}
\end{equation*}
$$

where $\langle\cdot, \cdot\rangle$ and $\langle\cdot, \cdot\rangle_{\Gamma_{N}}$ denote duality pairing on $\Omega$ and $\Gamma_{N}$, respectively.
The analysis for Neumann boundary and distributed optimal control problems was studied in [28].

Proposition 2.2 (Modification of Proposition 2.3 in [28]). For every $g_{N} \in$ $L^{2}\left(\Gamma_{N}\right)$ and $Q \in L^{2}(\Omega)$ with $g_{D}=0$, the Boussinesq equations (2.10)-(2.12) have a soultion $(\boldsymbol{u}, T, p) \in \boldsymbol{V} \times H^{1}(\Omega) \times L_{0}^{2}(\Omega)$. Moreover, if $(\boldsymbol{u}, T, p)$ is a solution to (2.10)-(2.12), then $(\boldsymbol{u}, T, p) \in \boldsymbol{V} \cap \boldsymbol{H}^{2}(\Omega) \times H^{s}(\Omega) \times L_{0}^{2}(\Omega) \cap H^{1}(\Omega)\left(1 \leq s<\frac{3}{2}\right)$ and there is a continuous function $P_{s}$ for each $s$ such that

$$
\begin{equation*}
\|\boldsymbol{u}\|_{H^{2}(\Omega)}+\|p\|_{H^{1}(\Omega)}+\|T\|_{H^{s}(\Omega)} \leq P_{s}\left(\|Q\|_{L^{2}(\Omega)}+\left\|g_{N}\right\|_{L^{2}\left(\Gamma_{N}\right)}\right) \tag{2.14}
\end{equation*}
$$

We describe the optimal control problem involving the functional (1.6) and state the optimality system. In the same way, one can study the optimal control problems for the functional (1.7) with the same distributed and Neumann boundary temperature control.

We look for a $\left(\boldsymbol{u}, T, p, Q, g_{N}\right) \in \boldsymbol{H}_{0}^{1}(\Omega) \times H^{1}(\Omega) \times L_{0}^{2}(\Omega) \times L^{2}(\Omega) \times \mathcal{V}$ such that the cost functional

$$
\begin{equation*}
\mathcal{J}_{1}\left(\boldsymbol{u}, T, p, Q, g_{N}\right)=\frac{1}{2 \alpha} \int_{\Omega}|\nabla \times \boldsymbol{u}|^{2} d \boldsymbol{x}+\frac{\beta}{2} \int_{\Omega}|Q|^{2} d \boldsymbol{x}+\frac{\gamma}{2} \int_{\Gamma_{N}}\left|g_{N}\right|^{2} d s \tag{2.15}
\end{equation*}
$$

is minimized subject to the constraints

$$
\begin{gather*}
\operatorname{Pr} a_{0}(\boldsymbol{u}, \boldsymbol{v})+c_{0}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})+b(\boldsymbol{v}, p)=d(T, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega),  \tag{2.16}\\
b(\boldsymbol{u}, q)=0 \quad \forall q \in L_{0}^{2}(\Omega),  \tag{2.17}\\
a_{1}(T, S)+c_{1}(\boldsymbol{u}, T, S)=\langle Q, S\rangle-\left\langle g_{N}, S\right\rangle_{\Gamma_{N}} \quad \forall S \in H_{D}^{1}(\Omega),  \tag{2.18}\\
T=g_{D} \quad \text { on } \Gamma_{1}, \quad T=0 \quad \text { on } \Gamma_{\mathrm{top}}, \tag{2.19}
\end{gather*}
$$

where $\mathcal{V}$ is a nonempty, closed and convex subset of $L^{2}\left(\Gamma_{N}\right)$.

Using the Lagrange multiplier method, the optimality system can be obtained as follows: find $(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in \boldsymbol{H}_{0}^{1}(\Omega) \times H^{1}(\Omega) \times L_{0}^{2}(\Omega) \times \boldsymbol{H}_{0}^{1}(\Omega) \times H_{D}^{1}(\Omega) \times$ $L_{0}^{2}(\Omega)$ such that

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}(\boldsymbol{u}, \boldsymbol{v})+c_{0}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})+b(\boldsymbol{v}, p)=d(T, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega),  \tag{2.20}\\
b(\boldsymbol{u}, q)=0 \quad \forall q \in L_{0}^{2}(\Omega), \\
a_{1}(T, S)+c_{1}(\boldsymbol{u}, T, S)=-\left\langle\frac{\Phi}{\beta}, S\right\rangle-\left\langle\frac{\Phi}{\gamma}, S\right\rangle_{\Gamma_{N}} \quad \forall S \in H_{D}^{1}(\Omega), \\
T=g_{D} \quad \text { on } \Gamma_{1}, \quad T=0 \quad \text { on } \Gamma_{\text {top }}, \\
\operatorname{Pr} a_{0}(\boldsymbol{\theta}, \boldsymbol{\xi})+c_{0}(\boldsymbol{\theta}, \boldsymbol{u}, \boldsymbol{\xi})+c_{0}(\boldsymbol{u}, \boldsymbol{\theta}, \boldsymbol{\xi})+b(\boldsymbol{\theta}, \sigma) \\
=\frac{1}{\alpha}(\nabla \times \boldsymbol{u}, \nabla \times \boldsymbol{\theta})-c_{1}(\boldsymbol{\theta}, T, \Phi) \quad \forall \boldsymbol{\theta} \in \boldsymbol{H}^{1}(\Omega), \\
b(\boldsymbol{\xi}, r)=0 \quad \forall r \in L_{0}^{2}(\Omega), \\
a_{1}(\Phi, \varphi)-c_{1}(\boldsymbol{u}, \varphi, \Phi)=d(\varphi, \boldsymbol{\xi}) \quad \forall \varphi \in H_{D}^{1}(\Omega)
\end{array}\right.
$$

### 2.2. Finite element approximation and error estimates

In this section we investigate a finite element discretization of the optimality system and the error estimates of the approximation solutions. First we choose a family of the finite dimensional subspaces $\boldsymbol{V}^{h} \subset \boldsymbol{H}^{1}(\Omega), V^{h} \subset H^{1}(\Omega), O^{h} \subset L^{2}(\Omega)$. We let $\boldsymbol{V}_{0}^{h}=\boldsymbol{V}^{h} \cap \boldsymbol{H}_{0}^{1}(\Omega), V_{D}^{h}=V^{h} \cap H_{D}^{1}(\Omega)$ and $O_{0}^{h}=O^{h} \cap L_{0}^{2}(\Omega)$. These families are parameterized by a parameter $h$ that tends to zero; commonly, $h$ is chosen to be some measure of the grid size. These finite-dimensional function spaces are defined on an approximate domain $\Omega_{h}$. For simplicity we will state our results in this section by assuming $\Omega_{h}=\Omega$. We assume that these finite element spaces satisfy the following approximation properties: there exist an integer $k$ and a constant $C$, independent of $h$, such that

$$
\begin{align*}
& \inf _{\boldsymbol{v}^{h} \in \boldsymbol{V}^{h}}\left\|\boldsymbol{v}-\boldsymbol{v}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)} \leq C h^{m}\|\boldsymbol{v}\|_{\boldsymbol{H}^{m+1}(\Omega)} \quad \forall \boldsymbol{v} \in \boldsymbol{H}^{m+1}(\Omega), 1 \leq m \leq k,  \tag{2.21}\\
& \inf _{q^{h} \in O^{h}}\left\|q-q^{h}\right\|_{L^{2}(\Omega)} \leq C h^{m}\|q\|_{H^{m}(\Omega)} \quad \forall q \in H^{m}(\Omega), 1 \leq m \leq k,  \tag{2.22}\\
& \inf _{T^{h} \in V^{h}}\left\|T-T^{h}\right\|_{H^{1}(\Omega)} \leq C h^{m}\|T\|_{H^{m+1}(\Omega)} \quad \forall T \in H^{m+1}(\Omega), 1 \leq m \leq k \tag{2.23}
\end{align*}
$$

Here we may choose any pair of subspaces $\boldsymbol{V}^{h}, V^{h}$ and $O^{h}$ such that $\boldsymbol{V}_{0}^{h}, V_{D}^{h}$ and $O_{0}^{h}$ can be used for finding finite element approximations of solutions of Boussinesq equations. Thus we make the following standard assumptions, which are exactly those employed in well-known finite element methods for the Navier-Stokes equations and the energy equation. Next, we assume the inf-sup condition: there exists a constant $C$, independent of $h$, such that

$$
\inf _{0 \neq q^{h} \in O_{0}^{h}} \sup _{0 \neq \boldsymbol{v}^{b} \in \boldsymbol{V}_{0}^{h}} \frac{b\left(\boldsymbol{v}^{h}, q^{h}\right)}{\left\|\boldsymbol{v}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}\left\|q^{h}\right\|_{L^{2}(\Omega)}} \geq C .
$$

This condition assures the stability of finite element discretizations of the Navier-Stokes equations and also that of the optimality system (2.20). The references $[8,16,17,32]$ may also be consulted for a catalogue of finite element subspaces that meet the requirements of the above approximation properties and the inf-sup condition. Once the approximating subspaces have been chosen, we seek $\left(\boldsymbol{u}^{h}, T^{h}, p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \sigma^{h}\right) \in \boldsymbol{V}_{0}^{h} \times V^{h} \times O_{0}^{h} \times \boldsymbol{V}_{0}^{h} \times V_{D}^{h} \times O_{0}^{h}$ by solving the discrete optimality system of equations:

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}\left(\boldsymbol{u}^{h}, \boldsymbol{v}^{h}\right)+c_{0}\left(\boldsymbol{u}^{h}, \boldsymbol{u}^{h}, \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, p^{h}\right)=d\left(T^{h}, \boldsymbol{v}^{h}\right) \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h},  \tag{2.24}\\
b\left(\boldsymbol{u}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in O_{0}^{h}, \\
a_{1}\left(T^{h}, S^{h}\right)+c_{1}\left(\boldsymbol{u}^{h}, T^{h}, S^{h}\right)=-\left\langle\frac{\Phi^{h}}{\beta}, S^{h}\right\rangle-\left\langle\frac{\Phi^{h}}{\gamma}, S^{h}\right\rangle_{\Gamma_{N}} \forall S^{h} \in V_{D}^{h}, \\
T^{h}=g_{D} \quad \text { on } \Gamma_{1}, \quad T^{h}=0 \quad \text { on } \Gamma_{\text {top }}, \\
\operatorname{Pr} a_{0}\left(\boldsymbol{\theta}^{h}, \boldsymbol{\xi}^{h}\right)+c_{0}\left(\boldsymbol{\theta}^{h}, \boldsymbol{u}^{h}, \boldsymbol{\xi}^{h}\right)+c_{0}\left(\boldsymbol{u}^{h}, \boldsymbol{\theta}^{h}, \boldsymbol{\xi}^{h}\right)+b\left(\boldsymbol{w}^{h}, \sigma^{h}\right) \\
=\frac{1}{\alpha}\left(\nabla \times \boldsymbol{u}^{h}, \nabla \times \boldsymbol{\theta}^{h}\right)-c_{1}\left(\boldsymbol{\theta}^{h}, T^{h}, \Phi^{h}\right) \quad \forall \boldsymbol{\theta}^{h} \in \boldsymbol{V}^{h}, \\
b\left(\boldsymbol{\xi}^{h}, r^{h}\right)=0 \quad \forall r^{h} \in O_{0}^{h}, \\
a_{1}\left(\Phi^{h}, S^{h}\right)-c_{1}\left(\boldsymbol{u}^{h}, S^{h}, \Phi^{h}\right)=d\left(S^{h}, \boldsymbol{\xi}^{h}\right) \quad \forall S^{h} \in V_{D}^{h}
\end{array}\right.
$$

We concern ourselves with questions related to the accuracy of finite element approximations in this section. The error estimate makes use of the results of [7] and [20] concerning the approximation of a class of nonlinear problems.

Here for the sake of completeness, we will state the relevant results specialized to our needs. The nonlinear problems considered in [7] and [16] are of the type

$$
\begin{equation*}
F(\lambda, \psi) \equiv \psi+L G(\lambda, \psi)=0 \tag{2.25}
\end{equation*}
$$

where $L \in \mathcal{L}(Y ; X), G$ is a $C^{2}$ mapping from $\Lambda \times X$ into $Y$, where $X$ and $Y$ are Banach spaces and $\Lambda$ is a compact interval of $\mathbb{R}$. We say that $\{(\lambda, \psi(\lambda)) \mid \lambda \in \Lambda\}$ is a branch of solutions of (2.25) if $\lambda \rightarrow \psi(\lambda)$ is a continuous function from $\Lambda$ into $X$ such that $F(\lambda, \psi(\lambda))=0$. The branch is called a nonsingular branch if we also have that $D_{\psi} F(\lambda, \psi(\lambda))$ is an isomorphism from $X$ into $X$ for all $\lambda \in \Lambda$. Here, $D_{\psi}$ denotes the Frechet derivative with respect to $\psi$. Approximations are defined by introducing a subspace $X^{h} \subset X$ and an approximating operator $L^{h} \in \mathcal{L}\left(Y ; X^{h}\right)$. Then we seek $\psi^{h} \in X^{h}$ such that

$$
\begin{equation*}
F^{h}\left(\lambda, \psi^{h}\right) \equiv \psi^{h}+L^{h} G\left(\lambda, \psi^{h}\right)=0 \tag{2.26}
\end{equation*}
$$

We will assume that there exists another Banach space $Z$, contained in $Y$, with continuous imbedding such that

$$
\begin{equation*}
D_{\psi} G(\lambda, \psi) \in \mathcal{L}(X ; Z), \quad \forall \lambda \in \Lambda, \quad \forall \psi \in X \tag{2.27}
\end{equation*}
$$

Concerning the operator $L^{h}$, we assume the approximation properties

$$
\begin{equation*}
\lim _{h \rightarrow 0}\left\|\left(L^{h}-L\right) y\right\|_{X}=0 \quad \forall y \in Y \tag{2.28}
\end{equation*}
$$

and

$$
\begin{equation*}
\lim _{h \rightarrow 0}\left\|L^{h}-L\right\|_{\mathcal{L}(Z ; X)}=0 \tag{2.29}
\end{equation*}
$$

Note that (2.27) and (2.29) imply that the operator $D_{\psi} G(\lambda, \psi) \in \mathcal{L}(X, X)$ is compact. Moreover, (2.29) follows from (2.28) whenever the imbedding $Z \subset Y$ is compact.

Now we can state the first result of [7] and [16] that used in the sequel.
Theorem 2.3. Let $X$ and $Y$ be Banach spaces and $\Lambda$ a compact subset of $\mathbb{R}$. Assume that $G$ is a $C^{2}$ mapping from $\Lambda \times X$ into $Y$ and that $D^{2} G$ is bounded on all sets of $\Lambda \times X\left(D^{2} G\right.$ represents second Fréchet derivative of $\left.G\right)$. Assume that (2.27)-(2.29) hold and $\{(\lambda, \psi(\lambda)) \mid \lambda \in \Lambda\}$ is a branch of nonsingular solutions of (2.25). Then, there exists a neighborhood $\mathcal{O}$ of the origin in $X$ and for $h \leq h_{0}$ small enough, a unique $C^{2}$ function $\lambda \in \Lambda \rightarrow \psi^{h}(\lambda) \in X^{h}$ such that $\left\{\left(\lambda, \psi^{h}(\lambda)\right) \mid \lambda \in \Lambda\right\}$ is a branch of nonsingular solutions of (2.26) and $\psi^{h}(\lambda)-\psi(\lambda) \in \mathcal{O}$ for all $\lambda$. Moreover, there exists a constant $C>0$, independent of $h$ and $\lambda$, such that

$$
\begin{equation*}
\left\|\psi^{h}(\lambda)-\psi(\lambda)\right\|_{X} \leq C\left\|\left(L^{h}-L\right) G(\lambda, \psi(\lambda))\right\|_{X} \quad \forall \lambda \in \Lambda \tag{2.30}
\end{equation*}
$$

For the second result, we have to introduce two other Banach spaces $H$ and $W$, such that $W \subset X \subset H$, with continuous imbeddings and assume that
for all $w \in W$ the operator $D_{\psi} G(\lambda, w)$ may be extended as a linear operator of $\mathcal{L}(H ; Y)$,
and the mapping $w \rightarrow D_{\psi} G(\lambda, w)$ is continuous from $W$ onto $\mathcal{L}(H ; Y)$.
We also suppose that

$$
\begin{equation*}
\lim _{h \rightarrow 0}\left\|L^{h}-L\right\|_{\mathcal{L}(Y ; H)}=0 \tag{2.32}
\end{equation*}
$$

Then we may state the following additional result.
Theorem 2.4. Assume the hypotheses of Theorem 2.3 and also assume that (2.31) and (2.32) hold. Assume in addition that

$$
\begin{align*}
& \text { for each } \lambda \in \Lambda, \psi(\lambda) \in W \text { and the function } \\
& \lambda \rightarrow \psi(\lambda) \text { is continuous from } \Lambda \text { into } W \tag{2.33}
\end{align*}
$$

and

$$
\begin{equation*}
\text { for each } \lambda \in \Lambda, D_{\psi} F(\lambda, \psi(\lambda)) \text { is an isomorphism of } H \text {. } \tag{2.34}
\end{equation*}
$$

Then, for $h \leq h_{1}$, sufficiently small, there exists a constant $C$, independent of $h$ and $\lambda$, such that

$$
\begin{equation*}
\left\|\psi^{h}(\lambda)-\psi(\lambda)\right\|_{H} \leq C\left\|\left(L^{h}-L\right) G(\lambda, \psi(\lambda))\right\|_{H}+\left\|\psi^{h}(\lambda)-\psi(\lambda)\right\|_{X} \quad \forall \lambda \in \Lambda \tag{2.35}
\end{equation*}
$$

We begin by recasting the optimality system (2.20) and its discretization (2.24) into a form that fits into the frame work. Let $\lambda=\frac{1}{P r}$ and let

$$
\begin{aligned}
X & =\boldsymbol{H}_{0}^{1}(\Omega) \times H^{1}(\Omega) \times L_{0}^{2}(\Omega) \times \boldsymbol{H}_{0}^{1}(\Omega) \times H_{D}^{1}(\Omega) \times L_{0}^{2}(\Omega), \\
Y & =\boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega) \times \boldsymbol{H}^{-1}(\Omega) \times H^{-1}(\Omega), \\
Z & =\boldsymbol{L}^{\frac{3}{2}}(\Omega) \times L^{\frac{3}{2}}(\Omega) \times \boldsymbol{L}^{\frac{3}{2}}(\Omega) \times L^{\frac{3}{2}}(\Omega), \\
X^{h} & =\boldsymbol{V}_{0}^{h} \times V^{h} \times O_{0}^{h} \times \boldsymbol{V}_{0}^{h} \times V^{h} \times O_{0}^{h} .
\end{aligned}
$$

Note that $Z \subset Y$ with a compact imbedding. The operator $L \in \mathcal{L}(Y ; X)$ is defined by

$$
L(\Xi, \Theta, \eta, P)=(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in X
$$

where, for $(\Xi, \Theta, \eta, \rho)$ in $Y,(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma)$ in $X$ is defined by the solution of the following system:

$$
\left\{\begin{array}{l}
a_{0}(\boldsymbol{u}, \boldsymbol{v})+b(\boldsymbol{v}, p)=\langle\boldsymbol{\Xi}, \boldsymbol{v}\rangle \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega),  \tag{2.36}\\
b(\boldsymbol{u}, q)=0 \quad \forall q \in L_{0}^{2}(\Omega) \\
a_{1}(T, S)=\langle\Theta, S\rangle \quad \forall S \in H_{D}^{1}(\Omega) \\
T=g_{D} \quad \text { on } \Gamma_{1}, \quad T=0 \quad \text { on } \Gamma_{\mathrm{top}}, \\
a_{0}(\boldsymbol{v}, \boldsymbol{\xi})+b(\boldsymbol{v}, \sigma)=\langle\eta, \boldsymbol{v}\rangle \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega), \\
b(\boldsymbol{\xi}, q)=0 \quad \forall q \in L_{0}^{2}(\Omega) \\
a_{1}(\Phi, S)=\langle P, S\rangle \quad \forall S \in H_{D}^{1}(\Omega)
\end{array}\right.
$$

Note that the system is weakly coupled. Analogously, the operator $L^{h} \in \mathcal{L}\left(Y ; X^{h}\right)$ is defined by

$$
L^{h}(\Xi, \Theta, \eta, P)=\left(\boldsymbol{u}^{h}, T^{h}, p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \sigma^{h}\right) \in X^{h}
$$

where $\left(\boldsymbol{u}^{h}, T^{h}, p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \sigma^{h}\right)$ in $X^{h}$ is defined by the solutions of the following system:

$$
\left\{\begin{array}{l}
a_{0}\left(\boldsymbol{u}^{h}, \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, p^{h}\right)=\left\langle\Xi, \boldsymbol{v}^{h}\right\rangle \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h}  \tag{2.37}\\
b\left(\boldsymbol{u}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in S_{0}^{h} \\
a_{1}\left(T^{h}, S^{h}\right)=\left\langle\Theta, S^{h}\right\rangle \quad \forall S^{h} \in V_{D}^{h} \\
T^{h}=g_{D} \quad \text { on } \Gamma_{1}, \quad T^{h}=0 \quad \text { on } \Gamma_{\mathrm{top}}, \\
a_{0}\left(\boldsymbol{v}^{h}, \boldsymbol{\xi}^{h}\right)+b\left(\boldsymbol{v}^{h}, \sigma^{h}\right)=\left\langle\eta, \boldsymbol{v}^{h}\right\rangle \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h}, \\
b\left(\boldsymbol{\xi}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in S_{0}^{h} \\
a_{1}\left(\Phi^{h}, S^{h}\right)=\left\langle P, S^{h}\right\rangle \quad \forall S^{h} \in V_{D}^{h}
\end{array}\right.
$$

This system is weakly coupled in the same sense as the system (2.36).

Let $\Lambda$ denote a compact subset of $\mathbb{R}$. Next we define the nonlinear mapping $G: \Lambda \times X \rightarrow Y$ as follows: $G(\lambda,(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma))=(\Xi, \Theta, \eta, P)$ for $\lambda \in \Lambda$, $(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in X$ where $(\Xi, \Theta, \eta, P) \in Y$ is defined as follows:

$$
\left\{\begin{align*}
\langle\Xi, \boldsymbol{v}\rangle= & \lambda c_{0}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})-\lambda d(T, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega),  \tag{2.38}\\
\langle\Theta, S\rangle= & c_{1}(\boldsymbol{u}, T, S)-\left\langle\frac{\Phi}{\beta}, S\right\rangle+\left\langle\frac{\Phi}{\gamma}, S\right\rangle_{\Gamma_{N}} \quad \forall S \in H_{D}^{1}(\Omega), \\
\langle\eta, \boldsymbol{v}\rangle= & \lambda c_{0}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{\xi})+\lambda c_{0}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{\xi})-\lambda c_{1}(\boldsymbol{v}, T, \Phi) \\
& -\frac{\lambda}{\alpha}(\nabla \times(\nabla \times \boldsymbol{u}), \boldsymbol{v}) \quad \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega), \\
\langle P, S\rangle= & c_{1}(\boldsymbol{u}, S, \Phi)-d(S, \boldsymbol{\xi}) \quad \forall S \in H_{D}^{1}(\Omega) .
\end{align*}\right.
$$

It is easily seen that the optimality system (2.20) is equivalent to

$$
\begin{equation*}
(\boldsymbol{u}, T, \lambda p, \boldsymbol{\xi}, \Phi, \lambda \sigma)+L G(\lambda,(\boldsymbol{u}, T, \lambda p, \boldsymbol{\xi}, \Phi, \lambda \sigma))=0 \tag{2.39}
\end{equation*}
$$

and that the discrete optimality system (2.24) is equivalent to

$$
\begin{equation*}
\left(\boldsymbol{u}^{h}, T^{h}, \lambda p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \lambda \sigma^{h}\right)+L^{h} G\left(\lambda,\left(\boldsymbol{u}^{h}, T^{h}, \lambda p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \lambda \sigma^{h}\right)\right)=0 \tag{2.40}
\end{equation*}
$$

Thus we have recast our continuous and discrete optimality problems into a form that enables us to apply Theorems 2.3 and 2.4.

Proposition 2.5. The problem (2.36) has a unique solution belonging to $X$. Assume that (2.21)-(2.23) hold. Then, the problem (2.37) has a unique solution belonging to $X^{h}$. Let $(\tilde{\boldsymbol{u}}, \tilde{T}, \tilde{p}, \tilde{\boldsymbol{\xi}}, \tilde{\Phi}, \tilde{\sigma})$ and $\left(\tilde{\boldsymbol{u}}^{h}, \tilde{T}^{h}, \tilde{p}^{h}, \tilde{\boldsymbol{\xi}}^{h}, \tilde{\Phi}^{h}, \tilde{\sigma}^{h}\right)$ denotes the solution of (2.36) and (2.37), respectively. Then we also have that

$$
\begin{align*}
& \left\|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{p}-\tilde{p}^{h}\right\|_{L^{2}(\Omega)}+\left\|\tilde{T}-\tilde{T}^{h}\right\|_{H^{1}(\Omega)} \\
& +\left\|\tilde{\boldsymbol{\xi}}-\tilde{\boldsymbol{\xi}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{\sigma}-\tilde{\sigma}^{h}\right\|_{L^{2}(\Omega)}+\left\|\tilde{\Phi}-\tilde{\Phi}^{h}\right\|_{H^{1}(\Omega)} \rightarrow 0 \quad \text { as } h \rightarrow 0 \tag{2.41}
\end{align*}
$$

If, in addition, $(\tilde{\boldsymbol{u}}, \tilde{T}, \tilde{p}, \tilde{\boldsymbol{\xi}}, \tilde{\Phi}, \tilde{\sigma}) \in \boldsymbol{H}_{0}^{m+1}(\Omega) \times H^{m+1}(\Omega) \times H^{m} \cap L_{0}^{2}(\Omega) \times \boldsymbol{H}_{0}^{m+1}(\Omega) \times$ $H^{m+1}(\Omega) \times H^{m} \cap L_{0}^{2}(\Omega)$, then there exists a constant $C$, independent of $h$, such that

$$
\begin{aligned}
& \left\|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{p}-\tilde{p}^{h}\right\|_{L^{2}(\Omega)}+\left\|\tilde{T}-\tilde{T}^{h}\right\|_{H^{1}(\Omega)} \\
& +\left\|\tilde{\boldsymbol{\xi}}-\tilde{\boldsymbol{\xi}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{\sigma}-\tilde{\sigma}^{h}\right\|_{L^{2}(\Omega)}+\left\|\tilde{\Phi}-\tilde{\Phi}^{h}\right\|_{H^{1}(\Omega)} \\
& \leq C h^{m}\left(\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\tilde{p}\|_{H^{m}(\Omega)}+\|\tilde{T}\|_{H^{m+1}(\Omega)}\right. \\
& \left.\quad \quad+\|\tilde{\boldsymbol{\xi}}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\tilde{\phi}\|_{H^{m}(\Omega)}+\|\tilde{\theta}\|_{H^{m+1}(\Omega)}\right)
\end{aligned}
$$

Proof. First, it is well known [7] that the two Stokes problems in (2.36) have unique solutions ( $\tilde{\boldsymbol{u}}, \tilde{p})$ and $(\tilde{\boldsymbol{\xi}}, \tilde{\phi})$ belonging to $\boldsymbol{H}_{0}^{1}(\Omega) \times \boldsymbol{L}_{0}^{2}(\Omega)$, respectively. Also, the discrete Stokes problems in (2.37) have unique solutions ( $\tilde{\boldsymbol{u}}^{h}, \tilde{p}^{h}$ ) and $\left(\tilde{\boldsymbol{\xi}}^{h}, \tilde{\phi}^{h}\right)$ belonging to $\boldsymbol{V}_{0}^{h} \times O_{0}^{h}$, respectively. Moreover, we have that

$$
\left\|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{p}-\tilde{p}^{h}\right\|_{L^{2}(\Omega)} \rightarrow 0
$$

and

$$
\left\|\tilde{\boldsymbol{\xi}}-\tilde{\boldsymbol{\xi}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{\sigma}-\tilde{\sigma}^{h}\right\|_{L^{2}(\Omega)} \rightarrow 0
$$

as $h \rightarrow 0$, and if in addition, $(\tilde{\boldsymbol{u}}, \tilde{p}) \in \boldsymbol{H}_{0}^{m+1}(\Omega) \times H^{m} \cap L_{0}^{2}(\Omega)$ and $(\tilde{\boldsymbol{\xi}}, \tilde{\sigma}) \in$ $\boldsymbol{H}_{0}^{m+1}(\Omega) \times H^{m} \cap L_{0}^{2}(\Omega)$, we have that

$$
\left\|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{p}-\tilde{p}^{h}\right\|_{L^{2}(\Omega)} \leq C h^{m}\left(\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\tilde{p}\|_{H^{m}(\Omega)}\right)
$$

and

$$
\left\|\tilde{\boldsymbol{\xi}}-\tilde{\boldsymbol{\xi}}^{h}\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\tilde{\sigma}-\tilde{\sigma}^{h}\right\|_{L^{2}(\Omega)} \leq C h^{m}\left(\|\tilde{\boldsymbol{\xi}}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\tilde{\sigma}\|_{H^{m}(\Omega)}\right) .
$$

Next, it is also well known that the two second order elliptic problems (the third and last equations in (2.36)) have unique solutions $\tilde{T}$ and $\tilde{\Phi}$ belonging to $H^{1}(\Omega)$, respectively. From the Babuska's theory, the discrete second order elliptic problems (the third and last equations in (2.37)) have unique solutions $\tilde{T}^{h}$ and $\tilde{\Phi}^{h}$ belonging to $V^{h}$, respectively. Moreover, we have that

$$
\begin{equation*}
\left\|\tilde{T}-\tilde{T}^{h}\right\|_{H^{1}(\Omega)} \rightarrow 0 \quad \text { as } h \rightarrow 0 \tag{2.42}
\end{equation*}
$$

and

$$
\begin{equation*}
\left\|\tilde{\Phi}-\tilde{\Phi}^{h}\right\|_{H^{1}(\Omega)} \rightarrow 0 \quad \text { as } h \rightarrow 0 \tag{2.43}
\end{equation*}
$$

and if, in addition, $\tilde{T} \in H^{m+1}(\Omega)$ and $\tilde{\Phi} \in H^{m+1}(\Omega)$, we have that

$$
\begin{align*}
& \left\|\tilde{T}-\tilde{T}^{h}\right\|_{H^{1}(\Omega)} \leq C h^{m}\|\tilde{T}\|_{H^{m+1}(\Omega)}  \tag{2.44}\\
& \left\|\tilde{\Phi}-\tilde{\Phi}^{h}\right\|_{H^{1}(\Omega)} \leq C h^{m}\|\tilde{\Phi}\|_{H^{m+1}(\Omega)} \tag{2.45}
\end{align*}
$$

Theorem 2.6. Assume that $\Lambda$ is a compact interval of $\mathbb{R}$ and that there exists a branch $\{(\lambda, \psi(\lambda):=(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma)) \in \Lambda \times X\}$ of nonsingular solutions of the optimality system (2.20). Assume that the finite elements spaces $\boldsymbol{V}^{h}, \boldsymbol{O}^{h}$, $V^{h}, O^{h}$ satisfy the condition (2.21)-(2.23). Then, there exists a neighborhood $\mathcal{O}$ of the origin in $X$ and, for $h \leq h_{0}$, small enough, a unique branch $\left\{\left(\lambda, \psi^{h}(\lambda):=\right.\right.$
$\left.\left.\left(\boldsymbol{u}^{h}, T^{h}, p^{h}, \boldsymbol{\xi}^{h}, \Phi^{h}, \sigma^{h}\right)\right) \in \Lambda \times X^{h}\right\}$ of solutions of the discrete optimality system (2.20) such that $\psi^{h}(\lambda)-\psi(\lambda) \in \mathcal{O}$ for all $\lambda \in \Lambda$. Moreover,

$$
\begin{align*}
& \left\|\psi^{h}(\lambda)-\psi(\lambda)\right\|_{X} \\
& =\left\|\boldsymbol{u}(\lambda)-\boldsymbol{u}^{h}(\lambda)\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|p(\lambda)-p^{h}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|T(\lambda)-T^{h}(\lambda)\right\|_{H^{1}(\Omega)} \\
& \quad+\left\|\boldsymbol{\xi}(\lambda)-\boldsymbol{\xi}^{h}(\lambda)\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\sigma(\lambda)-\sigma^{h}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|\Phi(\lambda)-\Phi^{h}(\lambda)\right\|_{H^{1}(\Omega)} \rightarrow 0 \tag{2.46}
\end{align*}
$$

as $h \rightarrow 0$, uniformly in $\lambda \in \Lambda$. If, in addition, $(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in \boldsymbol{H}_{0}^{m+1}(\Omega) \times$ $H^{m+1}(\Omega) \times L_{0}^{2}(\Omega) \cap H^{m} \times \boldsymbol{H}_{0}^{m+1}(\Omega) \times H^{m+1}(\Omega) \times L_{0}^{2}(\Omega) \cap H^{m}$ for $\lambda \in \Lambda$, then there exists a constant $C$, independent of $h$, such that

$$
\begin{align*}
& \left\|\boldsymbol{u}(\lambda)-\boldsymbol{u}^{h}(\lambda)\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|p(\lambda)-p^{h}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|T(\lambda)-T^{h}(\lambda)\right\|_{H^{1}(\Omega)} \\
& +\left\|\boldsymbol{\xi}(\lambda)-\boldsymbol{\xi}^{h}(\lambda)\right\|_{\boldsymbol{H}^{1}(\Omega)}+\left\|\sigma(\lambda)-\sigma^{h}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|\Phi(\lambda)-\Phi^{h}(\lambda)\right\|_{H^{1}(\Omega)} \\
& \leq C h^{m}\left(\|\boldsymbol{u}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
& \left.\quad+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) \tag{2.47}
\end{align*}
$$

uniformly in $\lambda \in \Lambda$.
Proof. Clearly, $G$ is a $C^{\infty}$ polynomial map from $\Lambda \times X$ into $Y$. Therefore, using (2.1)-(2.3), (2.6), and (2.8), it is easily shown that $D^{2} G(\lambda, \cdot)$ is bounded on all bounded sets of $X$. Now, given $(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in X$, a direct computation yields that $(\tilde{\Xi}, \tilde{\Theta}, \tilde{\eta}, \tilde{P}) \in Y$ satisfies

$$
(\bar{\Xi}, \bar{\Theta}, \bar{\eta}, \bar{P})=D_{\psi} G(\lambda,(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma))(\overline{\boldsymbol{u}}, \bar{T}, \bar{p}, \overline{\boldsymbol{\xi}}, \bar{\Phi}, \bar{\sigma})
$$

for $(\overline{\boldsymbol{u}}, \bar{T}, \bar{p}, \bar{\xi}, \bar{\Phi}, \bar{\sigma}) \in X$ if and only if

$$
\begin{aligned}
\langle\overline{\boldsymbol{\Xi}}, \boldsymbol{v}\rangle= & \lambda c_{0}(\boldsymbol{u}, \overline{\boldsymbol{u}}, \boldsymbol{v})+\lambda c_{0}(\overline{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})-\lambda d(\bar{T}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega), \\
\langle\bar{\Theta}, S\rangle= & c_{1}(\overline{\boldsymbol{u}}, T, S)+c_{1}(\boldsymbol{u}, \bar{T}, S)-\left\langle\frac{\bar{\Phi}}{\beta}, S\right\rangle-\left\langle\frac{\bar{\Phi}}{\gamma}, S\right\rangle_{\Gamma_{N}} \quad \forall S \in H_{D}^{1}(\Omega), \\
\langle\bar{\eta}, \boldsymbol{v}\rangle= & \lambda c_{0}(\boldsymbol{v}, \overline{\boldsymbol{u}}, \boldsymbol{\xi})+\lambda c_{0}(\boldsymbol{v}, \boldsymbol{u}, \overline{\boldsymbol{\xi}})+\lambda c_{0}(\overline{\boldsymbol{u}}, \boldsymbol{v}, \boldsymbol{\xi})+\lambda c_{0}(\boldsymbol{u}, \boldsymbol{v}, \overline{\boldsymbol{\xi}}) \\
& -\lambda c_{1}(\boldsymbol{v}, \bar{T}, \Phi)-\lambda c_{1}(\boldsymbol{v}, T, \bar{\Phi})-\frac{\lambda}{\alpha}(\nabla \times(\nabla \times \boldsymbol{u}), \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_{0}^{1}(\Omega), \\
\langle\bar{P}, S\rangle= & c_{1}(\overline{\boldsymbol{u}}, S, \Phi)+c_{1}(\boldsymbol{u}, S, \bar{\Phi})-d(S, \overline{\boldsymbol{\xi}}) \quad \forall S \in H_{D}^{1}(\Omega) .
\end{aligned}
$$

Thus, it follows from (2.1)-(2.3), (2.6), and (2.8) that $D_{\psi} G(\lambda,(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma)) \in$ $\mathcal{L}(X, Y)$. On the other hand, since $(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma) \in X$ and $(\overline{\boldsymbol{u}}, \bar{T}, \bar{p}, \overline{\boldsymbol{\xi}}, \bar{\Phi}, \bar{\sigma}) \in X$, by the Sobolev imbedding theorem, $T, \Phi, \bar{T}$, and $\bar{\Phi} \in L^{6}(\Omega)$, and $\boldsymbol{u}, \boldsymbol{\xi}$, and $\overline{\boldsymbol{u}}, \overline{\boldsymbol{\xi}} \in \boldsymbol{L}^{6}(\Omega)$ and $\frac{\partial \boldsymbol{u}}{\partial x_{j}}, \frac{\partial \boldsymbol{v}}{\partial x_{j}}, \frac{\partial \bar{u}}{\partial x_{j}}, \frac{\partial \overline{\boldsymbol{v}}}{\partial x_{j}} \in \boldsymbol{L}^{2}(\Omega)$ for $j=1,2$. Then it follows that $(\bar{\Xi}, \bar{\eta}, \bar{Q}, \bar{\Theta}, \bar{P}) \in Z$ and that, for $(\boldsymbol{u}, T, p, t, \boldsymbol{\xi}, \Phi, \sigma) \in X, D_{\psi} G(\lambda,(\boldsymbol{u}, T, p, \boldsymbol{\xi}, \Phi, \sigma)) \in \mathcal{L}(X, Z)$. Next, we turn to the approximation properties of the operator $T$. From Proposition 2.5, we
have that (2.28) holds. Since the imbedding of $Z$ into $Y$ is compact, (2.29) follows from (2.28), and the (2.46) follows from Theorem 2.3. Also from Proposition 2.5, we may conclude that there exists a constant $C$, independent of $h$, such that

$$
\begin{aligned}
&\left\|\left(L-L^{h}\right) G(\lambda, \psi(\lambda))\right\|_{X} \leq C h^{m}\left(\|\boldsymbol{u}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p\|_{H^{m}(\Omega)}+\|T\|_{H^{m+1}}\right. \\
&\left.+\|\boldsymbol{\xi}\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma\|_{H^{m}(\Omega)}+\|\Phi\|_{H^{m+1}}\right)
\end{aligned}
$$

Then (2.47) follows from Theorem 2.3.
Now, we derive an estimate for the error of $\boldsymbol{u}^{h}$ and $\boldsymbol{\xi}^{h}, T^{h}, \Phi^{h}$ in the $L^{2}$-norm. Since $G(\lambda, \psi(\lambda))$ does not depend on $p$, or $\sigma$, we redefine $X=\boldsymbol{H}^{1}(\Omega) \times H^{1}(\Omega) \times$ $\boldsymbol{H}^{1}(\Omega) \times H^{1}(\Omega)$ and $X^{h}=\boldsymbol{V}^{h} \times V^{h} \times \boldsymbol{V}^{h} \times V^{h}, Y$ and $Z$ remain as before.

Theorem 2.7. Assume the hypotheses of Theorem 2.6. Then there exists a constant $C$, independent of $h$ such that

$$
\begin{align*}
& \left\|\boldsymbol{u}^{h}(\lambda)-\boldsymbol{u}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|T^{h}(\lambda)-T(\lambda)\right\|_{L^{2}(\Omega)} \\
& +\left\|\boldsymbol{\xi}^{h}(\lambda)-\boldsymbol{\xi}(\lambda)\right\|_{L^{2}(\Omega)}+\left\|\Phi^{h}(\lambda)-\Phi(\lambda)\right\|_{L^{2}(\Omega)} \\
& \leq C h^{m+1}\left(\|\boldsymbol{u}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
& \left.\quad+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) . \tag{2.48}
\end{align*}
$$

Proof. We must verify that (2.31)-(2.34) hold in our setting; then the approximation properties (2.21) and the results of Theorems 2.4 and 2.6 easily lead to the conclusion. In similar methods with [19], we can verify (2.31)-(2.34).

A consequence of Theorems 2.6 and 2.7 is the following corollary that gives estimates for the error in the approximation of the controls.

Corollary 2.8. Define the approximate control by

$$
\begin{aligned}
Q^{h} & =-\frac{\Phi^{h}}{\beta} \quad \text { in } \Omega \\
g_{N}^{h} & =-\frac{\Phi^{h}}{\gamma} \quad \text { on } \Gamma_{N}
\end{aligned}
$$

and assume the hypotheses of Theorem 2.6. Then, for $h \leq h_{0}$ sufficiently small, there exists a constant $C$, independent of $h$ such that

$$
\begin{align*}
&\left\|Q^{h}-Q\right\|_{H^{1}(\Omega)} \leq C h^{m}\left(\|\boldsymbol{u}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
&\left.+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) \tag{2.49}
\end{align*}
$$

and

$$
\begin{align*}
&\left\|g_{N}^{h}-g_{N}\right\|_{H^{\frac{1}{2}}\left(\Gamma_{N}\right)} \leq C h^{m}\left(\|\boldsymbol{u}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
&\left.+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) \tag{2.50}
\end{align*}
$$

In addition the hypotheses of Theorem 2.7 hold, then, for $h \leq h_{1}$ sufficiently small, there exists a constant $C$, independent of $h$ such that

$$
\begin{align*}
&\left\|Q^{h}-Q\right\|_{L^{2}(\Omega)} \leq C h^{m+1}\left(\|\boldsymbol{u}(\lambda)\|_{H^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
&\left.+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) \tag{2.51}
\end{align*}
$$

and

$$
\begin{align*}
\left\|g_{N}^{h}-g_{N}\right\|_{L^{2}\left(\Gamma_{N}\right)} \leq C h^{m+1 / 2} & \left(\|\boldsymbol{u}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|p(\lambda)\|_{H^{m}(\Omega)}+\|T(\lambda)\|_{H^{m+1}(\Omega)}\right. \\
& \left.+\|\boldsymbol{\xi}(\lambda)\|_{\boldsymbol{H}^{m+1}(\Omega)}+\|\sigma(\lambda)\|_{H^{m}(\Omega)}+\|\Phi(\lambda)\|_{H^{m+1}(\Omega)}\right) \tag{2.52}
\end{align*}
$$

Proof. Recall that $Q^{h}=-\Phi^{h} / \gamma$ in $\Omega$ and $g_{N}^{h}=-\Phi^{h} / \gamma$ on $\Gamma_{N}$. Then, (2.49) and (2.51) easily follow from (2.47) and (2.48), respectivley; (2.50) follows from (2.47) and the trace theorem (see [1]), i.e., $\left\|g_{N}^{h}-g_{N}\right\|_{H^{\frac{1}{2}}\left(\Gamma_{N}\right)} \leq C\left\|\Phi^{h}-\Phi\right\|_{H^{1}(\Omega)}$. Finally, (2.52) follows from (2.47), (2.48) and the well-known inequality $\left\|g_{N}^{h}-g_{N}\right\|_{L^{2}\left(\Gamma_{N}\right)} \leq$ $C\left[\epsilon\left\|\Phi^{h}-\Phi\right\|_{H^{1}(\Omega)}+(1 / \epsilon)\left\|\Phi^{h}-\Phi\right\|_{L^{2}(\Omega)}\right]$ with the choice $\epsilon=h^{1 / 2}$.

## 3. The gradient method

The discrete optimality system of equations (2.24) consists of three groups of equations: the state equations for $\left(\boldsymbol{u}^{h}, p^{h}, T^{h}\right)$, the adjoint state equations for $\left(\boldsymbol{\xi}^{h}, \Phi^{h}, \sigma^{h}\right)$, and the optimality conditions for $g^{h}$ and $Q^{h}$. We may construct an iterative method, i.e., to iterate among the three groups of equations so that at each iteration we are dealing with a smaller size system of equations. Actually $\left(\boldsymbol{u}^{h}(k), p^{h}(k)\right)$ and $\left(\boldsymbol{\xi}^{h}(k), \sigma^{h}(k)\right)$ are solved with $T^{h}(k)$ and $\Phi^{h}(k)$ computed from the heat equation with $\boldsymbol{u}^{h}(k-1)$ and $\boldsymbol{\xi}^{h}(k-1)$ at each state equations respectively.

In this section, we consider the case of vorticity minimization problem with a distributed control. The other cases, such as a vorticity minimization problem with a Neumann boundarry control and a matching problem with a distributed or Neumann boundary control, may be studied in the same manner.

The gradient method for minimizing a functional $\mathcal{T}\left(Q^{h}\right):=\mathcal{J}\left(\boldsymbol{u}\left(Q^{h}\right), T\left(Q^{h}\right), Q^{h}\right)$ on a Hilbert space $O^{h}$ may be described as follows:

Algorithm 3.1. Gradient algorithm

1. Set $k=0$ and choose $Q^{h}(0)$.
2. Set $\delta Q^{h}:=-R \frac{d \mathcal{T}\left(Q^{h}(k)\right)}{d Q}$, and compute $\rho(k)=\arg \min _{\rho>0} \mathcal{T}\left(Q^{h}(k)+\rho \delta Q^{h}\right)$.
3. Set $Q^{h}(k+1)=Q^{h}(k)+\frac{\rho(k)}{\beta} \delta Q^{h}$.
4. Set $k=k+1$ and go to 2 .

Here $R$ is the Riesz map from the dual space of $O^{h}$ to $O^{h}$ and $\rho(k)$ is a sequence of positive step lengths.

Using Fréchet derivative, one can obtain the optimality system and relations between controls and adjoint variables. Thus, by the similar argument one can have the following gradient method, refer to [22]. A gradient method for a vorticity minization problem with a distributed control with variable step lengths $\{\rho(k)\}$ is given as follows. The Riesz map on $O^{h}$ can be taken as $I^{-1}$, i.e., the identity map.
(1) Choose an initial guess $Q^{h}(0)$.
(2) For each $k \geq 1$,
(a) solve for $\left(\boldsymbol{u}^{h}(k), T^{h}(k), p^{h}(k)\right)$ from the state equation with $Q^{h}(k-1)$, for $\forall S^{h} \in V^{h}, \forall v^{h} \in V_{0}^{h}$ and $\forall q^{h} \in O_{0}^{h}$,

$$
\left\{\begin{array}{l}
a_{1}\left(T^{h}(k), S^{h}\right)+c_{1}\left(\boldsymbol{u}^{h}(k-1), T^{h}(k), S^{h}\right)=\left\langle Q^{h}(k-1), S^{h}\right\rangle  \tag{3.1}\\
T^{h}(k)=g_{D} \quad \text { on } \Gamma_{1}, \quad T^{h}(k)=0 \quad \text { on } \Gamma_{\mathrm{top}}, \\
\operatorname{Pr} a_{0}\left(\boldsymbol{u}^{h}(k), \boldsymbol{v}^{h}\right)+c_{0}\left(\boldsymbol{u}^{h}(k), \boldsymbol{u}^{h}(k), \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, p^{h}\right)-d\left(T^{h}(k), \boldsymbol{v}^{h}\right)=0, \\
b\left(\boldsymbol{u}^{h}(k), q^{h}\right)=0
\end{array}\right.
$$

(b) solve $\left(\xi^{h}(k), \Phi^{h}(k), \sigma^{h}(k)\right)$ from the adjoint state equation with $\left(\boldsymbol{u}^{h}(k), T^{h}(k), p^{h}(k)\right), \forall \varphi^{h} \in V_{D}^{h}, \forall \boldsymbol{\theta}^{h} \in \boldsymbol{V}_{0}^{h}$ and $\forall r^{h} \in O_{0}^{h}$,

$$
\left\{\begin{array}{l}
a_{1}\left(\Phi^{h}(k), \varphi^{h}\right)+c_{1}\left(\boldsymbol{u}^{h}(k), \varphi^{h}, \Phi^{h}(k)\right)=d\left(\varphi^{h}, \boldsymbol{\xi}^{h}(k-1)\right),  \tag{3.2}\\
\Phi^{h}(k)=0 \text { on } \Gamma_{D}, \\
\operatorname{Pr} a_{0}\left(\boldsymbol{\theta}^{h}, \boldsymbol{\xi}^{h}(k)\right)+c_{0}\left(\boldsymbol{\theta}^{h}, \boldsymbol{u}^{h}(k), \boldsymbol{\xi}^{h}(k)\right)+c_{0}\left(\boldsymbol{u}^{h}(k), \boldsymbol{\theta}^{h}, \boldsymbol{\xi}^{h}(k)\right) \\
+b\left(\boldsymbol{\theta}^{h}, \sigma^{h}(k)\right)=\frac{1}{\alpha}\left(\nabla \times \boldsymbol{u}^{h}(k), \nabla \times \boldsymbol{\theta}^{h}\right)-c_{1}\left(\boldsymbol{\theta}^{h}, T^{h}(k), \Phi^{h}(k)\right), \\
b\left(\boldsymbol{\xi}^{h}(k), r^{h}\right)=0 .
\end{array}\right.
$$

(c) solve for $Q^{h}(k)$ from the optimality condition

$$
\begin{equation*}
Q^{h}(k)=(1-\rho(k)) Q^{h}(k-1)-\frac{\rho(k)}{\beta} \Phi^{h}(k) \quad \text { on } \Omega \tag{3.3}
\end{equation*}
$$

The convergence property of the above algorithm is given in the following result. The convergence of the conjugate gradient algorithm can be proved in the similar fashion.

Theorem 3.1. Let $\left(\boldsymbol{u}^{h}(k), p^{h}(k), T^{h}(k), \boldsymbol{\xi}^{h}(k), \psi^{h}(k), \Phi^{h}(k), Q^{h}(k)\right)$ be the solution of the gradient algorithm and ( $\boldsymbol{u}^{h}, p^{h}, T^{h}, \boldsymbol{\xi}^{h}, \psi^{h}, \Phi^{h}, Q^{h}$ ) the solution of (2.24) with $\left\langle\frac{\Phi^{h}}{\gamma}, S^{h}\right\rangle_{\Gamma_{N}}=0 \forall S^{h} \in V_{D}^{h}$. Then if $\rho(k)>0$, is sufficiently small, $Q^{h}(k) \rightarrow$ $Q^{h}$ and thus, $\left(\boldsymbol{u}^{h}(k), p^{h}(k), T^{h}(k), \boldsymbol{\xi}^{h}(k), \psi^{h}(k), \Phi^{h}(k)\right) \rightarrow\left(\boldsymbol{u}^{h}, p^{h}, T^{h}, \boldsymbol{\xi}^{h}, \psi^{h}, \Phi^{h}\right)$ as $k \rightarrow \infty$.

Proof. We will make use of the following classical result; see, e.g., [8]. Let $\mathcal{X}$ be a Hilbert space with norm $\|\cdot\|_{\mathcal{X}}$ and scalar product $(\cdot, \cdot)_{\mathcal{X}}$. Let $\mathcal{T}(\cdot)$ be a real valued function on $\mathcal{X}$. Suppose that $\mathcal{T}(\cdot)$ is of class $C^{2}$, that it has a local minimum at $Q \in \mathcal{X}$ and that there exist two real numbers $M, n$ and a ball $B \subset \mathcal{X}$ centered at $Q$ such that for all $\tilde{Q} \in B$ and all $\delta Q_{1}, \delta Q_{2} \in \mathcal{X}$ we have that

$$
\begin{equation*}
\mathcal{T}^{\prime \prime}(\tilde{Q}) \cdot\left(\delta Q_{1}, \delta Q_{2}\right) \leq M\left\|\delta Q_{1}\right\|_{X}\left\|\delta Q_{2}\right\|_{X} \quad \text { and } \quad m\left\|\delta Q_{1}\right\|_{X}^{2} \leq \mathcal{T}^{\prime \prime}(\tilde{Q}) \cdot\left(\delta Q_{1}, \delta Q_{1}\right) \tag{3.4}
\end{equation*}
$$

where $\mathcal{T}^{\prime \prime}(\tilde{Q}) \cdot\left(\delta Q_{1}, \delta Q_{2}\right)$ is the bilinear form associated with the second Fréchet derivatives of $\mathcal{T}(\cdot)$. Suppose that $\rho(k)$ is chosen so that

$$
\begin{equation*}
0<\rho_{*} \leq \rho(k) \leq \rho^{*}<\frac{2 m}{M^{2}} \quad \text { for all } k \tag{3.5}
\end{equation*}
$$

for some positive numbers $\rho_{*}$ and $\rho^{*}$. Then the iterates of the algorithm

$$
\begin{equation*}
Q^{h}(k)=Q^{h}(k-1)-\rho(k) \mathcal{T}\left(Q^{h}(k-1)\right), \quad n=1,2, \ldots, \tag{3.6}
\end{equation*}
$$

converges to $Q^{h}$ for any initial guess $Q^{h}(0) \in B$.
Now, let $\mathcal{X}=O^{h}$. For each $\tilde{Q}^{h} \in O^{h}$, the second Fréchet derivative $\mathcal{T}^{\prime \prime}\left(\tilde{Q}^{h}\right)$. ( $\delta Q_{1}^{h}, \delta Q_{2}^{h}$ ) may be computed by

$$
\begin{equation*}
\mathcal{T}^{\prime \prime}\left(\tilde{Q}^{h}\right) \cdot\left(\delta Q_{1}^{h}, \delta Q_{2}^{h}\right)=\beta\left(\delta Q_{1}^{h}, \delta Q_{2}^{h}\right)+\frac{1}{\alpha}\left(\nabla \times \tilde{\boldsymbol{w}}_{1}^{h}, \nabla \times \tilde{\boldsymbol{w}}_{2}^{h}\right)+\frac{1}{\alpha}\left(\nabla \times \tilde{\boldsymbol{u}}^{h}, \nabla \times \tilde{\boldsymbol{z}}^{h}\right) \tag{3.7}
\end{equation*}
$$

where $\tilde{\boldsymbol{u}}^{h} \in \boldsymbol{V}_{0}^{h}$ is the solution of

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}\left(\tilde{\boldsymbol{u}}^{h}, \boldsymbol{v}^{h}\right)+c_{0}\left(\tilde{\boldsymbol{u}}^{h}, \tilde{\boldsymbol{u}}^{h}, \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, \tilde{p}^{h}\right)=d\left(\tilde{T}^{h}, \boldsymbol{v}^{h}\right) \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h}  \tag{3.8}\\
b\left(\tilde{\boldsymbol{u}}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in O_{0}^{h} \\
a_{1}\left(\tilde{T}^{h}, S^{h}\right)+c_{1}\left(\tilde{\boldsymbol{u}}^{h}, \tilde{T}^{h}, S^{h}\right)=\left\langle\tilde{Q}^{h}, S^{h}\right\rangle \quad \forall S^{h} \in V_{D}^{h} \\
\tilde{T}^{h}=g_{D} \quad \text { on } \Gamma_{1}, \quad \tilde{T}^{h}=0 \quad \text { on } \Gamma_{\text {top }}
\end{array}\right.
$$

the first variations $\tilde{\boldsymbol{w}}_{1}^{h}$ and $\tilde{\boldsymbol{w}}_{2}^{h}$ are solutions of

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}\left(\tilde{\boldsymbol{w}}_{i}^{h}, \boldsymbol{v}^{h}\right)+c_{0}\left(\tilde{\boldsymbol{w}}_{i}^{h}, \tilde{\boldsymbol{u}}^{h}, \boldsymbol{v}^{h}\right)  \tag{3.9}\\
+c_{0}\left(\tilde{\boldsymbol{u}}^{h}, \tilde{\boldsymbol{w}}_{i}^{h}, \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, \tilde{r}_{i}^{h}\right)=d\left(\tilde{A}_{i}^{h}, \boldsymbol{v}^{h}\right) \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h}, \\
b\left(\tilde{\boldsymbol{w}}_{i}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in O_{0}^{h} \\
a_{1}\left(\tilde{A}_{i}^{h}, S^{h}\right)+c_{1}\left(\tilde{\boldsymbol{w}}_{i}^{h}, \tilde{T}^{h}, S^{h}\right)+c_{1}\left(\tilde{\boldsymbol{u}}^{h}, \tilde{A}_{i}^{h}, S^{h}\right)=\left\langle\delta Q_{i}^{h}, S^{h}\right\rangle \quad \forall S^{h} \in V_{D}^{h}, \\
\tilde{A}_{i}^{h}=0 \quad \text { on } \Gamma_{D}
\end{array}\right.
$$

for $i=1,2$, respectively, the second variation $\tilde{\boldsymbol{z}}^{h}$ is the solution of

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}\left(\tilde{\boldsymbol{z}}^{h}, \boldsymbol{v}^{h}\right)+c_{0}\left(\tilde{\boldsymbol{z}}^{h}, \tilde{\boldsymbol{u}}^{h}, \boldsymbol{v}^{h}\right)  \tag{3.10}\\
+c_{0}\left(\tilde{\boldsymbol{w}}_{1}^{h}, \tilde{\boldsymbol{w}}_{2}^{h}, \boldsymbol{v}^{h}\right)+b\left(\boldsymbol{v}^{h}, \tilde{s}^{h}\right)=d\left(\tilde{B}^{h}, \boldsymbol{v}^{h}\right) \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h}, \\
b\left(\tilde{\boldsymbol{z}}^{h}, q^{h}\right)=0 \quad \forall q^{h} \in O_{0}^{h}, \\
a_{1}\left(\tilde{B}^{h}, S^{h}\right)+c_{1}\left(\tilde{\boldsymbol{z}}^{h}, \tilde{T}^{h}, S^{h}\right) \\
+c_{1}\left(\tilde{\boldsymbol{w}}_{1}^{h}, \tilde{A}_{2}^{h}, S^{h}\right)+c_{1}\left(\tilde{\boldsymbol{w}}_{2}^{h}, \tilde{A}_{1}^{h}, S^{h}\right)=0 \quad \forall S^{h} \in V_{D}^{h}, \\
\tilde{B}^{h}=0 \quad \text { on } \Gamma_{D} .
\end{array}\right.
$$

Let $\hat{Q}^{h}$ be the initial guess in the ball of radius $\zeta$, i.e., $\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|=\rho \leq \zeta$. Now, we need to show that there exists a $\zeta$ such that $\forall \rho \leq \zeta$, (3.4) is satisfied for some $M$ and $m$.

From (2.14), we have with continuous function $f_{1}$ and a constant $K$ that

$$
\begin{equation*}
\left\|\tilde{\boldsymbol{u}}^{h}\right\|_{1}+\left\|\tilde{T}^{h}\right\|_{1} \leq f_{1}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|+K_{1}\right) \tag{3.11}
\end{equation*}
$$

By the same way in the proof of Proposition 2.1, we have that there is a continuous function $f_{2}$ and $f_{3}$ such that

$$
\begin{equation*}
\left\|\tilde{\boldsymbol{w}}_{i}^{h}\right\|_{1} \leq f_{2}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right)\left\|\delta Q_{i}^{h}\right\| \tag{3.12}
\end{equation*}
$$

for $i=1,2$, and

$$
\begin{equation*}
\left\|\tilde{\boldsymbol{z}}^{h}\right\|_{1} \leq f_{3}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right)\left\|\delta Q_{1}^{h}\right\|\left\|\delta Q_{2}^{h}\right\| \tag{3.13}
\end{equation*}
$$

for $i=1,2$. Then using (3.11), (3.12) and (3.13) in (3.4), we have that, for some constant $K_{2}>0$,

$$
\begin{align*}
& \left.\left.\mid \mathcal{T}^{\prime \prime}\left(\tilde{\boldsymbol{u}}^{h}\left(\tilde{Q}^{h}\right)\right), \tilde{Q}^{h}\right)\right) \cdot\left(\delta Q_{1}^{h}, \delta Q_{2}^{h}\right) \mid \\
& \leq\left(\beta+\frac{K_{2}}{\alpha} f_{2}^{2}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right)+\frac{K_{2}}{\alpha}\left(f_{1}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|+K_{1}\right) f_{3}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right)\right)\right. \\
& \quad \times\left\|\delta Q_{1}^{h}\right\|\left\|\delta Q_{2}^{h}\right\| \tag{3.14}
\end{align*}
$$

and

$$
\begin{align*}
& \left.\left.\mid \mathcal{T}^{\prime \prime}\left(\tilde{\boldsymbol{u}}^{h}\left(\tilde{Q}^{h}\right)\right), \tilde{Q}^{h}\right)\right) \cdot\left(\delta Q_{1}^{h}, \delta Q_{1}^{h}\right) \mid \\
& \geq\left(\beta-\frac{K_{2}}{\alpha}\left(f_{1}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|+K_{1}\right) f_{3}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right)\right)\left\|\delta Q_{1}^{h}\right\|^{2}\right. \tag{3.15}
\end{align*}
$$

Now, from the continuity of $f_{1}, f_{2}$ and $f_{3}$ we have that for every positive $\epsilon$ there exists a $\zeta$ such that if $\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|<\zeta$ then $f_{i}\left(\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|\right) \leq \epsilon$ for $i=1,2,3$. Thus,

$$
\begin{equation*}
\left|\mathcal{T}^{\prime \prime}\left(\tilde{Q}^{h}\right) \cdot\left(\delta Q_{1}^{h}, \delta Q_{2}^{h}\right)\right| \leq\left(\beta+\frac{K_{2}}{\alpha}\left(2 \epsilon+K_{1}\right) \epsilon\right)\left\|\delta Q_{1}^{h}\right\|\left\|\delta Q_{2}^{h}\right\|=M\left\|\delta Q_{1}^{h}\right\|\left\|\delta Q_{2}^{h}\right\| \tag{3.16}
\end{equation*}
$$

and

$$
\begin{equation*}
\left|\mathcal{T}^{\prime \prime}\left(\tilde{Q}^{h}\right) \cdot\left(\delta Q_{1}^{h}, \delta Q_{1}^{h}\right)\right| \geq\left(\beta-\frac{K_{2}}{\alpha}\left(\epsilon+K_{1}\right) \epsilon\right)\left\|\delta Q_{1}^{h}\right\|^{2}=m\left\|\delta Q_{1}^{h}\right\|^{2}, \tag{3.17}
\end{equation*}
$$

where

$$
M=\beta+\frac{K_{2}}{\alpha}\left(2 \epsilon+K_{1}\right) \epsilon \quad \text { and } \quad m=\beta-\frac{K_{2}}{\alpha}\left(\epsilon+K_{1}\right) \epsilon .
$$

If $\epsilon$ is chosen such that $\beta-\frac{K_{2}}{\alpha}\left(\epsilon+K_{1}\right) \epsilon \geq 0$ then there exists a $\zeta$ such that $\left\|\tilde{Q}^{h}-\hat{Q}^{h}\right\|=\rho \leq \zeta$, (3.16) and (3.17) imply the inequalities in (3.4).

Now we discuss the numerical solution of the optimal control problem. To carry out the computation we discretized the problem using the finite element method. We use the Taylor-Hood finite element, that is, the piecewise quadratic element for the velocity and the temperature and piecewise linear element for the pressure defined on a triangle mesh.

In the equations (3.1), the Navier-Stokes equations are coupled with heat equation. Since we want to use the existent tcodes for Navier-Stokes equations and heat equation, we use the block Gauss-Seidel iteration. Equations (3.2) are also solved using the block Gauss-Seidel iteration.

Since the Navier-Stokes equations in (3.1) are nonlinear, we use Newton's method based on exact Jacobian. At each Newton's iteration, we solve the linear system of equations by Gaussian eliminations for sparse matrices. Since quadratic convergence of Newton's method is valid only within a contraction ball, we normally first perform a few (usually 3 or 4 times) simple successive iterations and then switch to the Newton's method. The simple successive iterations are defined by

$$
\left\{\begin{array}{l}
\operatorname{Pr} a_{0}\left(\boldsymbol{u}^{h}(k), \boldsymbol{v}^{h}\right)+c_{0}\left(\boldsymbol{u}^{h}(k-1), \boldsymbol{u}^{h}(k), \boldsymbol{v}^{h}\right) \\
+b\left(\boldsymbol{v}^{h}, p^{h}\right)-d\left(T^{h}(k), \boldsymbol{v}^{h}\right)=0 \quad \forall \boldsymbol{v}^{h} \in \boldsymbol{V}_{0}^{h} \\
b\left(\boldsymbol{u}^{h}(k), q^{h}\right)=0 \quad \forall q^{h} \in O_{0}^{h}
\end{array}\right.
$$

In the case of the uncontrolled Navier-Stokes equations, the solution is unique for a small Reynolds number and the simple successive approximations converges globally and linearly (See [17]).

In the same way, one can study the optimal control problems for the functional (1.7) with the same distributed and Neumann boundary temperature control.

To decide $\rho(k)$, one can solve a optimization problem 2 in Algorithm 3.1. Since this optimization problem is nonlinear, one can approximate this problem to a linearized optimization problem using Taylor series. Also, one can let $\rho=1$ and choose the prescribed tolerance $\tau$ and perform the following steps in each iteration $k=1, \ldots$,
(1) if $\mathcal{T}\left(Q^{h}(k)\right) \geq \mathcal{T}\left(Q^{h}(k-1)\right)$, set $\rho(k)=l \times \rho(k-1)$ and go to the beginning of the iterations; otherwise, go to next step;
(2) if $\left|\mathcal{T}\left(Q^{h}(k)\right)-\mathcal{T}\left(Q^{h}(k-1)\right)\right| /\left|\mathcal{T}\left(Q^{h}(k)\right)\right|>\tau$, set $\rho(k)=m \times \rho(k-1)$ and go to beginning of the iterations; otherwise, stop where $l<1$ and $m>1$.
In our problem, we find that the convergence is extremely senstive with $\rho(k)$. The above algorithm is used to find an approximation of $\rho(k)$. Then we fix the step length $\rho / \beta$ for each problem.

## 4. Computational results

In this section we test two examples involving the functionals (1.6)-(1.7) with distributed and Neumann boundary controls using the gradient algorithm studied in Section 3. Let us consider that the domain $\Omega$ is the unit square $(0,1) \times(0,1) \subset \mathbb{R}^{2}$.

The state variables satisfy the following Boussinesq equations:

$$
\left\{\begin{array}{l}
-\Delta \boldsymbol{u}+(\boldsymbol{u} \cdot \nabla) \boldsymbol{u}+\nabla p=9.8 T \boldsymbol{j} \quad \text { in } \Omega  \tag{4.1}\\
\nabla \cdot \boldsymbol{u}=0 \quad \text { in } \Omega \\
-\Delta T+(\boldsymbol{u} \cdot \nabla) T=Q \quad \text { in } \Omega
\end{array}\right.
$$

In our computation, we take the Prandtl number to be 1 . We also take $\alpha=1$.
Since the distributed control and the Neumann control are equally effective in our computational experiments, we are not going to consider the case of both controls are applied. Thus, $\beta=0$ in the case of Neumann boundary control problem and, of course, $\gamma=0$ in the distributed control problem. The stopping criterion for the gradient method is as follows: if

$$
\frac{\left|\mathcal{T}\left(Q^{h}(k)\right)-\mathcal{T}\left(Q^{h}(k-1)\right)\right|}{\left|\mathcal{T}\left(Q^{h}(k)\right)\right|}<t o l,
$$

then stop for the distributed control problem and if

$$
\frac{\left|\mathcal{T}\left(g_{N}^{h}(k)\right)-\mathcal{T}\left(g_{N}^{h}(k-1)\right)\right|}{\left|\mathcal{T}\left(g_{N}^{h}(k)\right)\right|}<t o l,
$$

then stop for the Neumann boundary control problem. We set tol $=10^{-6}$ for all tests in this section.

To stabilize the convection flow, we test both vorticity minimization and matching problems.

### 4.1. Vorticity minimization

First, let us consider the vorticity minimization problem. Minimize the following functional

$$
\begin{equation*}
\mathcal{J}(\boldsymbol{u}, T, p, Q)=\frac{1}{2} \int_{\Omega}|\nabla \times \boldsymbol{u}|^{2} d \boldsymbol{x}+\frac{\beta}{2} \int_{\Omega}|Q|^{2} d \boldsymbol{x} \tag{4.2}
\end{equation*}
$$

for the case of distributed control or

$$
\begin{equation*}
\mathcal{J}(\boldsymbol{u}, T, p, g)=\frac{1}{2} \int_{\Omega}|\nabla \times \boldsymbol{u}|^{2} d \boldsymbol{x}+\frac{\gamma}{2} \int_{\Gamma_{N}}\left|g_{N}\right|^{2} d s \tag{4.3}
\end{equation*}
$$

for the Neumann boundary control subject to (4.1) with boundary conditions

$$
\begin{aligned}
\boldsymbol{u} & =\mathbf{0} \quad \text { on } \partial \Omega, \\
\frac{\partial T}{\partial n}=0 \quad \text { on } \Gamma_{\text {left }} \cup \Gamma_{\text {right }}, & T=0 \quad \text { on } \Gamma_{\text {top }}, \quad T=10 \quad \text { on } \Gamma_{1}, \\
\frac{\partial T}{\partial n} & =g_{N} \quad \text { on } \Gamma_{N},
\end{aligned}
$$

where $\Gamma_{\text {left }}=\{(0, y) \mid y \in(0,1)\}, \Gamma_{\text {right }}=\{(1, y) \mid y \in(0,1)\}, \Gamma_{1}=\{(x, 0) \mid x \in(0,0.2)\}$, $\Gamma_{\text {top }}=\{(x, 1) \mid x \in(0,1)\}, \Gamma_{N}=\{(x, 0) \mid x \in(0.2,1)\}$, see Fig. 1. When $g_{N}=0, Q=0$, we say that this problem is the uncontrolled problem. The numerical solution of the uncontrolled problem is shown in Fig. 2. Our uncontrolled system is generated by only a Dirichlet boundary condition and thus optimal states can not be easily achieved from distributed and/or Neumann boundary controls in geneal.


Fig. 2. Uncontrolled temperature field (left) and velocity field (right) ( $Q=0$ and $\left.g_{N}=0\right)$.

In a distributed control problem, the control is the distributed temperature source $Q$. Of course, we set $g_{N}=0$ in this case. Table 1 shows that the vorticity of $\boldsymbol{u}$ decrease as the value of $\beta$ goes to 0 . We choose the size of step $\frac{\rho}{\beta}=100$ in the simple gradient method. The value of $\|\nabla \times \boldsymbol{u}\|$ is reduced $96.0 \%$ when $\beta=10^{-5}$. The controlled temperature field and velocity field are plotted in Fig. 3.


Fig. 3. For vorticity minimization with distributed control, controlled temperature field (left) and velocity field (right), $\beta=10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$ from the top.

Table 1. Costs for vorticity minimization with distributed control, time step $\left(:=\frac{\rho}{\beta}\right)=100$.

| $\beta$ | $\\|\nabla \times \boldsymbol{u}\\|$ | $\\|Q\\|$ | $\mathcal{J}(\boldsymbol{u}, p, T, Q)$ |
| :---: | :---: | :---: | :---: |
| Uncontrolled | $0.4291090153 \mathrm{E}+00$ | $0.0000000000 \mathrm{E}+00$ | $0.9206727351 \mathrm{E}-01$ |
| $10^{-2}$ | $0.3351268557 \mathrm{E}+00$ | $0.1862048373 \mathrm{E}+01$ | $0.7349112541 \mathrm{E}-01$ |
| $10^{-3}$ | $0.1392299614 \mathrm{E}+00$ | $0.5719345338 \mathrm{E}+01$ | $0.2604794663 \mathrm{E}-01$ |
| $10^{-4}$ | $0.5279644311 \mathrm{E}-01$ | $0.8873134440 \mathrm{E}+01$ | $0.5330357942 \mathrm{E}-02$ |
| $10^{-5}$ | $0.1651178240 \mathrm{E}-01$ | $0.1176825243 \mathrm{E}+02$ | $0.8287783047 \mathrm{E}-03$ |

Table 2. Costs for vorticity minimization, with Neumann control, time step $\left(:=\frac{\rho}{\gamma}\right)=50$.

| $\gamma$ | $\\|\nabla \times \boldsymbol{u}\\|$ | $\\|g\\|$ | $\mathcal{J}(\boldsymbol{u}, p, T, g)$ |
| :---: | :---: | :---: | :---: |
| Uncontrolled | $0.4291090153 \mathrm{E}+00$ | $0.0000000000 \mathrm{E}+00$ | $0.9206727351 \mathrm{E}-01$ |
| $10^{-2}$ | $0.3265086000 \mathrm{E}+00$ | $0.2284611847 \mathrm{E}+01$ | $0.7940118941 \mathrm{E}-01$ |
| $10^{-3}$ | $0.1033243147 \mathrm{E}+00$ | $0.6676982954 \mathrm{E}+01$ | $0.2762900769 \mathrm{E}-01$ |
| $10^{-4}$ | $0.1797341718 \mathrm{E}-01$ | $0.8418176486 \mathrm{E}+01$ | $0.3704806630 \mathrm{E}-02$ |
| $10^{-5}$ | $0.2187092979 \mathrm{E}-02$ | $0.8837571617 \mathrm{E}+01$ | $0.3929050483 \mathrm{E}-03$ |

Table 3. CPU time in seconds (with number of iterations).

| $\beta$ or $\gamma$ | Distributed control | Neumann control |  |
| :---: | :---: | ---: | :---: |
| $10^{-2}$ | $32.117(16)$ | $24.475(14)$ |  |
| $10^{-3}$ | $97.067(42)$ | $117.399(71)$ |  |
| $10^{-4}$ | $278.649(166)$ | $213.765(126)$ |  |
| $10^{-5}$ | $714.049(412)$ | $323.254(214)$ |  |

In a Neumann boundary control problem, the control is the heat flux $g_{N}$ along the boundary $\Gamma_{N}$. Of course, we set $Q=0$ in this case. Table 2 shows that the vorticity of $\boldsymbol{u}$ decrease as the value of $\gamma$ goes to 0 . We choose the size of step $\frac{\rho}{\gamma}=50$ in the simple gradient method.

The value of $\|\nabla \times \boldsymbol{u}\|$ is reduced $99.0 \%$ when $\gamma=10^{-5}$. The controlled temperature field and velocity field are plotted in Fig. 4.

Table 3 shows CPU time in seconds for vorticity minimization problem. We use MacBook Pro having 2.16 GHz Intel Core 2 Duo CPU and Intel Fortran compiler V10.


Fig. 4. For vorticity minimization with Neumann control, controlled temperature field (left) and velocity field (right), $\gamma=10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$ from the top.

### 4.2. Stablizaiton via velocity matching problem

Let us consider the velocity matching problem. Minimize the following functional

$$
\begin{equation*}
\mathcal{J}(\boldsymbol{u}, T, p, Q)=\frac{1}{2} \int_{\Omega}\left|\boldsymbol{u}-\boldsymbol{U}_{d}\right|^{2} d \boldsymbol{x}+\frac{\beta}{2} \int_{\Omega}|Q|^{2} d \boldsymbol{x} \tag{4.4}
\end{equation*}
$$

for the case of distributed control or

$$
\begin{equation*}
\mathcal{J}\left(\boldsymbol{u}, T, p, g_{N}\right)=\frac{1}{2} \int_{\Omega}\left|\boldsymbol{u}-\boldsymbol{U}_{d}\right|^{2} d \boldsymbol{x}+\frac{\gamma}{2} \int_{\Gamma_{N}}\left|g_{N}\right|^{2} d s \tag{4.5}
\end{equation*}
$$

for the Neumann boundary control subject to (4.1) with boundary conditions

$$
\begin{aligned}
& \boldsymbol{u}=\mathbf{0} \quad \text { on } \partial \Omega, \\
& \frac{\partial T}{\partial n}=0 \quad \text { on } \Gamma_{\text {left }} \cup \Gamma_{\text {right }}, \quad T=0 \quad \text { on } \Gamma_{\text {top }}, \quad T=10 \quad \text { on } \Gamma_{1}, \\
& \frac{\partial T}{\partial n}=g_{N} \quad \text { on } \Gamma_{N} .
\end{aligned}
$$

In this case, we set $\boldsymbol{U}_{d}=0$ which is the special case of temperature tracking problem. Table 4 shows that the value of $\|\boldsymbol{u}\|$ decreases as the value of $\beta$ goes to 0 . We choose the size of step $\frac{\rho}{\beta}=10000$ in the simple gradient method. The value of $\|\boldsymbol{u}\|$ is reduced $97.0 \%$ when $\beta=10^{-7}$. The controlled temperature field and velocity field are plotted in Fig. 5.

Table 4. Costs for velcocity matching problem with distributed control, time step $\left(:=\frac{\rho}{\beta}\right)=10000$.

| $\beta$ | $\left\\|\boldsymbol{u}-\boldsymbol{U}_{d}\right\\|$ | $\\|Q\\|$ | $\mathcal{J}(\boldsymbol{u}, p, T, Q)$ |
| :---: | :---: | :---: | :---: |
| Uncontrolled | $0.5414021543 \mathrm{E}-01$ | $0.0000000000 \mathrm{E}+00$ | $0.1465581463 \mathrm{E}-02$ |
| $10^{-4}$ | $0.3537005264 \mathrm{E}-01$ | $0.2646210885 \mathrm{E}+01$ | $0.9756419143 \mathrm{E}-03$ |
| $10^{-5}$ | $0.1214415941 \mathrm{E}-01$ | $0.6071283264 \mathrm{E}+01$ | $0.2580427062 \mathrm{E}-03$ |
| $10^{-6}$ | $0.4763565779 \mathrm{E}-02$ | $0.8511036407 \mathrm{E}+01$ | $0.4756464983 \mathrm{E}-04$ |
| $10^{-7}$ | $0.1529331139 \mathrm{E}-02$ | $0.1095148889 \mathrm{E}+02$ | $0.7166182309 \mathrm{E}-05$ |

In a Neumann boundary control problem, the control is the heat flux $g_{N}$ along the boundary $\Gamma_{N}$. Of course, we set $Q=0$ in this case. Table 5 shows that the value of $\|\boldsymbol{u}\|$ decreases as the value of $\gamma$ goes to 0 . We choose the size of step $\frac{\rho}{\gamma}=5000$ in the simple gradient method.

The value of $\|\boldsymbol{u}\|$ is reduced $99.6 \%$ when $\gamma=10^{-5}$. The controlled temperature field and velocity field are plotted in Fig. 6.

Table 6 shows CPU time in seconds for velocity matching problem.
We should note that, as seen in Tables 3 and 6 , Neumann control is more effective than the distributed one from the computational points of view, especially when $\beta$ and $\gamma$ are relatively small.


Fig. 5. For velcocity matching problem with distributed control, controlled temperature field (left) and velocity field (right), $\beta=10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$ from the top.


Fig. 6. For velcocity matching problem with Neumann control, controlled temperature field (left) and velocity field (right), $\gamma=10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$ from the top.

Table 5. Costs for velcocity matching problem with Neumann control, time step $\left(:=\frac{\rho}{\gamma}\right)=5000$.

| $\gamma$ | $\left\\|\boldsymbol{u}-\boldsymbol{U}_{d}\right\\|$ | $\left\\|g_{N}\right\\|$ | $\mathcal{J}\left(\boldsymbol{u}, p, T, g_{N}\right)$ |
| :---: | :---: | :---: | :---: |
| Uncontrolled | $0.5414021543 \mathrm{E}-01$ | $0.0000000000 \mathrm{E}+00$ | $0.1465581463 \mathrm{E}-02$ |
| $10^{-4}$ | $0.3532286012 \mathrm{E}-01$ | $0.2920747642 \mathrm{E}+01$ | $0.1050390563 \mathrm{E}-02$ |
| $10^{-5}$ | $0.8904758594 \mathrm{E}-02$ | $0.7032880992 \mathrm{E}+01$ | $0.2869544380 \mathrm{E}-03$ |
| $10^{-6}$ | $0.1959647560 \mathrm{E}-02$ | $0.8355395057 \mathrm{E}+01$ | $0.3682642255 \mathrm{E}-04$ |
| $10^{-7}$ | $0.2288944106 \mathrm{E}-03$ | $0.8831095340 \mathrm{E}+01$ | $0.3925608571 \mathrm{E}-05$ |

Table 6. CPU time in seconds (with number of iterations).

| $\beta$ or $\gamma$ | Distributed control | Neumann control |  |
| :---: | :---: | ---: | :---: |
| $10^{-4}$ | $47.833(26)$ | $27.647(16)$ |  |
| $10^{-5}$ | $94.124(51)$ | $83.971(50)$ |  |
| $10^{-6}$ | $319.688(174)$ | $111.461(65)$ |  |
| $10^{-7}$ | $818.377(554)$ | $314.689(206)$ |  |

Acknowledgments. The author would like to thank the unknown referees for their numerous suggestions and comments and typographical and stylistic corrections, helping to improve presentation of the paper.

## References

[1] R. Adams, Sobolev Spaces. Academic, New York, 1975.
[2] F. Abergel and F. Casas, Some optimal control problems of multistate equations appearing in fluid mechanics. Math. Modelling Numer. Anal., 27 (1993), 223-247.
[3] W. Alt and K. Malanowski, The Lagrange-Newton method for state constrained optimal control problems. Comput. Optim. Appl., 4 (1995), 217-239.
[4] G.V. Alekseev, Solvability of stationary boundary control problems for heat convection equations. Siber. Math. Jour., 39 (1998), 844-858.
[5] F. Abergel and R. Temam, On Some Control Problems in Fluid Mechanics. Theoret. Comput. Fluid Dynamics, 1 (1990), 303-325.
[6] I. Babuska and A. Aziz, Survey lectures on the mathematical foundations of the finite element method. The Mathematical Foundations of the Finite Element Method with Applicaions to Partial Differential Equations, A. Aziz (ed.), Academic, New York, 1972, 3-359.
[7] F. Brezzi, J. Rappaz and P.A. Raviart, Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions. Numer. Math., 36 (1980), 1-25.
[ 8 ] P. Cialet, The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam, 1978.
[9] P. Cuvelier, Optimal control of a system governed by the Navier-Stokes equations coupled with the heat equations. New Developments in Differential Equations, W. Eckhaus (ed.), North-Holland, Amsterdam, 1976, 81-98.
[10] J.E. Dennis, M. Heinkenschloss and L.N. Vincente, Trust-region interior-point SQP algorithms for a class of nonlinear programming problems. SIAM J. Control Optim., 36 (1998), 1750-1794.
[11] M. Desai and K. Ito, Optimal controls of Navier-Stokes equations. SIAM J. Control Optim., 32 (1994), 1428-1446.
[12] H.O. Fattorini and S.S. Sritharan, Optimal controls for viscous flow problems. Proc. Roy. Soc. London Ser. A, 439 (1992), 81-102.
[13] A.V. Fursikov, Ptoperties of solutions to some extrimal problems related to the NavierStokes system. Mat. Sb., 118 (1982), 323-349.
[14] A.V. Fursikov and O.Yu. Imanuvilov, Local exact boundary controllability of the Boussinesq equation. SIAM J. Control Optim., 36 (1998), 391-421.
[15] O. Ghattas and J.J. Bark, Optimal control of two- and three-dimensional incompressible Navier-Stokes flows. J. Comput. Phys., 136 (1997), 231-244.
[16] V. Girault and P.-A. Raviart, Finite Element Methods for Navier-Stokes Equations. Springer, Berlin, 1986.
[17] M. Gunzburger, Finite Element Methods for Incompressible Viscous Flows: A Guide to Theory, Practice and Algorithms. Academic, Boston, MA, 1989.
[18] M. Gunzburger, L. Hou and T. Svobodny, Heating and cooling control of temperature distributions along boundaries of flow domains. J. Math. Syst. Estim. Control, 3 (1993), 147-172.
[19] M. Gunzburger, L. Hou and T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls. Math. Comput., 57 (1991), 123-125.
[20] M. Gunzburger, L. Hou and T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with Dirichlet controls. Math. Modelling Numer. Anal., 25 (1991), 711-748.
[21] M. Gunzburger, L. Hou and T. Svobodny, Boundary velocity control of incompressible flow with an application to viscous drag reduction. SIAM J. Control Optim., 30 (1992), 167-181.
[22] M. Gunzburger and H. Lee, Analysis, approximation, and computation of a coupled solid/fluid temperature control problem. Compt. Methods Appl. Mech. Engng., 118 (1994), 133-152.
[23] M. Heinkenschloss, Formulation and analysis of a sequential quadratic programming method for the optimal Dirichlet boundary control of Navier-Stokes flow. Optimal Control: Theory, Algorithms, and Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, 178-203.
[24] M. Hinze and K. Kunisch, Second order methods for optimal control of time-dependent fluid flow. SIAM J. Control Optim., 40 (2001), 925-946.
[25] L. Hou, Analysis and finite element approximation of some optimal control problem associated with the Navier-Stokes equations. Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, 1989.
[26] K. Ito and S.S. Ravindran, Optimal control of thermally convected fluid flows. SIAM J. Sci. Comput., 19 (1998), 1847-1869.
[27] H.-C. Lee and O.Yu. Imnuvilov, Analysis of Neumann boundary optimal control problems for the stationary Boussinesq equations including solid media. SIAM J. Control Optim., 457-477.
[28] H.-C. Lee and O.Yu. Imnuvilov, Analysis of optimal control problems for the 2-D stationary Boussinesq equations. J. Math. Anal. Appl., 242 (2000), 191-211.
[29] G. Savaré, Regularity and perturbation results for mixed second order elliptic problems. Commun. in P.D.E., 22 (1997), 869-899.
[30] G. Savaré, Elliptic equations in Lipshitz domains. J. Func. Anal., 152 (1998), 176-201.
[31] V. Tikhomirov, Fundamental Principles of the Theory of Extremal Problems. Wiley, Chichester, 1982.
[32] R. Temam, Navier-Stokes Equations. North-Holland, Amsterdam, 1979.
[33] C. Taylor and P. Hood, A numerical solution of the Navier-Stokes equations using the finite element methods. Comput. and Fluids, 1 (1973), 73-100.


[^0]:    This work was supported in part by grant No. R01-2006-000-10472-0 from the Korea Science and Engineering Foundation.

