Topology Optimization for Steady-state anisothermal flow targeting solid with piecewise constant thermal diffusivity

Alexandre Vieira, Alain Bastide, Pierre-Henri Cocquet

Noname manuscript No.
(will be inserted by the editor)

Topology optimization for steady-state anisothermal flow targeting solids with piecewise constant thermal diffusivity.

Alexandre Vieira · Alain Bastide ·
Pierre-Henri Cocquet

the date of receipt and acceptance should be inserted later

Abstract Several engineering problems result in a PDE-constrained optimization problem that aims at finding the shape of a solid inside a fluid which minimizes a given cost function. These problems are categorized as Topology Optimization (TO) problems. In order to tackle these problems, the solid may be located with a penalization term added in the constraints equations that vanishes in fluid regions and becomes large in solid regions. This paper addresses a TO problem for anisothermal flows modelled by the steady-state incompressible Navier-Stokes system coupled to an energy equation, with mixed boundary conditions, under the Boussinesq approximation. We first prove the existence and uniqueness of a solution to these equations as well as the convergence of its finite element discretization. Next, we show that our TO problem has at least one optimal solution for cost functions that satisfy general assumptions. The convergence of discrete optimum toward the continuous one is then proved as well as necessary first order optimality conditions. Eventually, all these results let us design a numerical algorithm to solve a TO problem approximating solids with piecewise constant thermal diffusivities also refered as multi-materials. A physical problem solved numerically for varying parameters concludes this paper.

Keywords Topology optimization; Navier-Stokes equations; Anisothermal flows; Multi-materials.

Mathematics Subject Classification (2010) 76D55 · 49Q10 · 93C20 · 65K99

Physique et Ingénierie Mathématique pour l'Énergie et l'Environnement (PIMENT), Université de la Réunion, 2 rue Joseph Wetzell, 97490 Sainte-Clotilde, France. (Alexandre.Vieira@univ-reunion.fr, Alain.Bastide@univ-reunion.fr) · Laboratoire des Sciences de l'Ingénieur Appliquées à la Mécanique et au Génie Electrique (SIAME), E2S-UPPA, Université de Pau et des Pays de l'Adour, 64000 Pau, France (Pierre-Henri.Cocquet@univ-reunion.fr,Pierre-Henri.Cocquet@univ-pau.fr)

1 Introduction

Finding the shape of a solid located inside a fluid that either minimizes or maximizes a given physical effect has several applications in engineering and applied sciences (see [44, 45, 48, 49] for several examples).

There exist various mathematical methods to deal with such problems that fall into the class of PDE-constrained optimization. The topological asymptotic expansion [6, 19, 46] considers the solid as a hole or an inhomogeneity with characteristic size ε. The so-called *topological gradient* is then defined as the first order term in the asymptotic expansion of the cost function as $\varepsilon \to 0$. The shape optimization method [31, 44, 45] computes the gradient of the cost function with respect to perturbation of the boundary of the solid also referred to as *shape derivative*. Once the gradient of the cost function is computed, the geometry of the computational domain changes and thus these two methods usually need some specific techniques to follow the evolution of the mesh while numerically solving the state equations.

In this paper, we choose to locate the solid thanks to a penalization term added in the Navier-Stokes equation. This term vanishes in the fluid zone and goes to infinity in a solid region of the computational domain [7]. This non-smooth binary function is usually replaced by a smooth approximation, referred as *interpolation function* [50]. This smooth approximation can then be used in gradient-based optimization algorithms. Using such model as constraint in the shape optimization problem is referred to as a *topology optimization* (TO) problem. It is worth noting that the major drawback of this approach is that the solid is only located when the velocity of the fluid is smaller that a given tolerance, thus producing some *grey* regions, while the topological expansion and the shape optimization methods produce *black and white* solutions, exactly locating the solid. Nevertheless, this approach does not need specific remeshing techniques.
We refer to the review papers [3, 26] for many references that deals with numerical resolution of TO problems applied to several different physical settings. More precisely, we refer to [2, 4, 17, 50, 51] for numerical and physical studies on TO involving heat transfers in fluid flows (involving for instance natural, forced or mixed convection) since this is going to be the physical setting of interest of this paper.

Regarding the mathematical analysis of TO problems using the penalization term, we first note that they amount to find some coefficients in the PDE that minimize a given criterion and can thus be seen as parametric optimization problems [9]. They also share similarities with problems that seek to recover some unknown parameters in the PDE from measurements [20, 40, 47]. However, all the aforementioned references deal with scalar coercive problem (see also [24] that deals with Helmholtz equation which is elliptic but not coercive) and, even if they give some insight on how to mathematically tackle a TO problem for fluid flows, they can not be used to study the problem of interest in this paper. We also refer to [36] for several results on discretization of a general PDE-constrained optimization problem. It is however worth noting that our TO problem does not fit in the framework of [36] since the constraints equations considered depend linearly on the control.

In the literature, the mathematical study of TO problems for fluid flows using penalization remains scarce. We refer to [14] where a TO problem for incompressible Stokes equations have been studied. In [30], some existence results as well as some limitations of the shape optimization using a penalization technique are given for the incompressible Stokes equation. We finally refer to [33] where a shape optimization problem combining perimeter regularization, penalization technique and phase-field approach have been introduced.

Considering the previous literature review, there is, to the best of our knowledge, no mathematical study (existence, approximation and convergence of optimal solution) of a TO problem involving heat transfers in anisothermal flows. In addition, in the TO mathematical literature [14, 30, 33], the boundary conditions considered are homogeneous Dirichlet on the whole boundary. This simplifies the mathematical analysis of the incompressible Navier-Stokes equation since the non-linear term vanishes after integrating by part hence simplifying the derivation of a priori estimates [13, 25, 32, 54]. The first objectives of this paper are then to study a general TO problem involving, as constraint, the Navier-Stokes equation coupled to the heat equation with mixed (homogeneous/inhomogeneous Dirichlet and traction) boundary conditions since the latter are closer to those used in physical situations [2, 50, 51].

Another topic of interest of this paper is to look for optimized solid with thermal conductivity that are not only constant as it is the case in most TO studies. There already exists some methods to get optimized physical parameter that are piecewise constant [41, 43, 55] but the latter introduce an optimization parameter per constant which may thus lead to large optimization problems. This constraint have been lifted [56] where an ordered SIMP (Solid Isotropic Material with Penalization) interpolation function for the elastic modulus is introduced. Although this technique could be applied to get optimized piecewise constant thermal diffusivity, its definition is actually based on *gluing* together power curves which results in a piecewise-defined function with some points where it is not differentiable (see [56, Figure 2]). Another goal of this paper is then to introduce a smooth globally defined interpolation function that yields an approximation of the optimized piecewise constant thermal parameters.

Plan of the paper

The paper is now organized as follows: first we introduce the PDE modelling heat transfers in anisothermal flows, namely the steady-state Navier-Stokes system under the Boussinesq assumption coupled to an energy equation. The fluid/solid interpolation function as well as the multi-material interpolation function that will be used to obtain optimized solid approximating a piecewise constant thermal diffusivity are introduced next. Then, we will end this introduction by clearly defining the optimization problem under study.

We then study the existence and uniqueness of a weak solution to the constraint equations. We prove next the convergence of a finite element approximation of the latter where a discrete optimization parameter is used. After giving some general conditions on the cost function to obtain the existence of an optimal solution, we then prove the convergence of the discrete optimum toward its continuous counterpart. We end this paper with numerical simulations to show the interest of our approach.

Definition of the topology optimization problem

We present now the main ideas leading to the TO problem considered in this paper. First, since the velocity of the fluid vanishes inside the solid, one can use a penalization model as introduced in [7] in order to write the fluid-solid model as a single system valid on the whole computational domain. Its solution converges toward the one of the fluid-solid interface problem. Such model involves an indicator function, that is a binary variable, to locate the solid and thus makes the optimization problem intractable [30]. To bypass this difficulty, some smooth regularization of the indicator function is introduced hence defining another model where the location of the solid now depends on a continuous variable. We emphasize that the approach described above has been used in several works [3, 26] and we describe it below for the case of anisothermal flows.

Let $\Omega' \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$. We assume the fluid occupies a region $\Omega'_f \subset \Omega'$ and that a solid is defined by a region Ω'_s such that $\Omega' = \Omega'_f \cup \Omega'_s$. The Boussinesq approximation (see e.g. [51] for the steady case) of the Navier-Stokes equation coupled to convective heat transfer reads:

$$\begin{aligned} \nabla \cdot \bar{u} &= 0 \quad \text{in } \Omega'_f, \\ (\bar{u} \cdot \nabla)\, \bar{u} + \frac{1}{\rho_0} \nabla \bar{p} - \nu \Delta \bar{u} - \beta g (T - T_0) e_y &= 0 \quad \text{in } \Omega'_f, \\ \rho_0 c_p \nabla \cdot (\bar{u} T) - \nabla \cdot \left(\bar{k}(\bar{x}) \nabla T \right) &= 0 \quad \text{in } \Omega', \\ \bar{u} &= 0 \quad \text{in } \Omega'_s, \end{aligned} \tag{1}$$

where $\bar{u} : \Omega' \to \mathbb{R}^d$ is the velocity vector, $\bar{p}$ (scalar) the pressure, T (scalar) the temperature, ρ_0 is the density at the reference temperature T_0, c_p is the heat capacity of the fluid, ν is the kinematic viscosity, g the gravity acting in direction $-e_y$, $\bar{k}$ is the spatially varying thermal conductivity.

In order to get the dimensionless form of (1), denote:

$$x = \frac{\bar{x}}{L_0}, \; u = \frac{\bar{u}}{V_0}, \; p = \frac{\bar{p}}{\rho_0 V_0^2}, \; \theta = \frac{T - T_0}{\delta T},$$

where L_0 and V_0 are reference length and velocity, and δT denotes a constant temperature. The equations now read:

$$\nabla \cdot u = 0 \qquad \text{in } \Omega_f, \tag{2a}$$

$$(u \cdot \nabla)\, u + \nabla p - \mathrm{Re}^{-1} \Delta u - \mathrm{Ri}\, \theta e_y = 0 \qquad \text{in } \Omega_f, \tag{2b}$$

$$\nabla \cdot (u\theta) - \nabla \cdot \left(\mathrm{Re}^{-1}\, \mathrm{Pr}^{-1}\, k(x) \nabla \theta \right) = 0 \qquad \text{in } \Omega, \tag{2c}$$

$$u = 0 \qquad \text{in } \Omega_s, \tag{2d}$$

where $\Omega = \Omega'/L$ (the same goes for Ω_f and Ω_s), $\mathrm{Re} = (V_0 L_0)/\nu$ is the Reynolds number, $\mathrm{Ri} = (g\beta L_0 \delta T)/V_0^2$ the Richardson number, $\mathrm{Pr} = \rho_0 \nu c_p / k_f$ is the Prandtl number, $k(x) = \bar{k}(L_0 x)/k_f$ is the (ratio of) thermal diffusivities.

As explained above, the solid can be located thanks to a penalization term [7] of the form $\eta^{-1} \mathbf{1}_{\Omega_s} u$ added in the momentum conservation equation (2b). The latter formally enforces, as $\eta \to 0$, that $u|_{\Omega_s} = 0$ as well as a no-slip boundary condition on $\partial\Omega_s$. For (2), this reads

$$\begin{aligned} \nabla \cdot u &= 0 \text{ in } \Omega, \\ (u \cdot \nabla)\, u + \nabla p - \mathrm{Re}^{-1} \Delta u - \mathrm{Ri}\theta e_y + \tfrac{1}{\eta} \mathbf{1}_{\Omega_s}(x) u &= 0 \text{ in } \Omega, \\ \nabla \cdot (u\theta) - \nabla \cdot \left(\mathrm{Re}^{-1} \mathrm{Pr}^{-1} k(x) \nabla\theta \right) &= 0 \text{ in } \Omega, \end{aligned} \tag{3}$$

where $k(x) = \frac{1}{k_f} \left(k_f \mathbf{1}_{\Omega_s}(x) + k_s(1 - \mathbf{1}_{\Omega_s}(x)) \right)$.

To get the model studied in this paper that acts as constraint in the TO problem, we introduce the function $\alpha : x \in \Omega \mapsto \mathbb{R}^+$ which is going to be a parameter locating the solid in Ω. We now consider some smooth regularization $h_\tau(\alpha(x))$ of the indicator function that satisfy

$$\begin{aligned} h_\tau(s) &\xrightarrow[\tau \to +\infty]{} 0 \text{ for } s < \alpha_0, \\ h_\tau(s) &\xrightarrow[\tau \to +\infty]{} \alpha_{\max} \text{ for } s \geq \alpha_0, \end{aligned}$$

and $\alpha_{\max} = \eta^{-1}$, where the convergence is pointwise. In addition, the fluid/solid zones can now be obtained as

$$\Omega_s := \{ x \in \Omega \mid \alpha(x) < \alpha_0 \}, \; \Omega_f := \{ x \in \Omega \mid \alpha(x) \geq \alpha_0 \}.$$

We emphasize that we can now use $\alpha : x \in \Omega \mapsto \alpha(x) \in [0, \alpha_{\max}]$ as a design parameter to locate the solid zones inside Ω.

The dimensionless form of the energy and penalized incompressible Navier-Stokes equations under Boussinesq assumption are finally written as follows

$$\nabla \cdot u = 0 \qquad \text{in } \Omega, \tag{4a}$$

$$(u \cdot \nabla)\, u = -\nabla p + A\Delta u - h_\tau(\alpha) u + B\theta e_y \qquad \text{in } \Omega, \tag{4b}$$

$$\nabla \cdot (u\theta) = \nabla \cdot (C k_\tau(\alpha) \nabla\theta) \qquad \text{in } \Omega, \tag{4c}$$

where A, B, C are physical constants that are introduced to lighten the overall expressions and reads

$$A = \frac{1}{\mathrm{Re}}, \; B = \mathrm{Ri}, \; C = \frac{1}{\mathrm{Re}\,\mathrm{Pr}}.$$

Regarding the boundary conditions, we are going to work with mixed boundary conditions and first assume $\partial\Omega = \Gamma$ is Lipschitz and can be decomposed as $\Gamma = \Gamma_{\mathrm{w}} \cup \Gamma_{\mathrm{in}} \cup \Gamma_{\mathrm{out}}$ with $|\Gamma_{\mathrm{w}}| > 0$, $|\Gamma_{\mathrm{in}}| > 0$, $|\Gamma_{\mathrm{out}}| > 0$ and $\overline{\Gamma_{\mathrm{in}}} \cap \overline{\Gamma_{\mathrm{out}}} = \emptyset$. Here, Γ_{w} are the walls, Γ_{in} the inlet/entrance and Γ_{out} is the exit/outlet of the computational domain. These boundary conditions [50, 51] read

$$u = u_{\mathrm{in}}, \qquad \theta = 0, \qquad \text{in } \Gamma_{\mathrm{in}}, \tag{5a}$$

$$u = 0, \qquad k_\tau \partial_n \theta = \phi, \qquad \text{in } \Gamma_{\mathrm{w}}, \tag{5b}$$

$$A\partial_n u - np = 0, \qquad \partial_n \theta = 0, \qquad \text{in } \Gamma_{\mathrm{out}}, \tag{5c}$$

where $u_{\mathrm{in}} \in H^{1/2}_{00}(\Gamma_{\mathrm{in}})^d$ is a given function (where $H^{1/2}_{00}(\Gamma_{\mathrm{in}})$ is the set of $\psi \in H^{1/2}(\Gamma_{\mathrm{in}})$ such that $\psi_{00} \in H^{1/2}(\Gamma)$ where ψ_{00} denotes the extension by zero of ψ on Γ), ϕ is a given heat flux, n is the unit normal vector to the boundary, and ∂_n is the normal derivative.

Definitions of the interpolation functions

All the results we will present in sections 2 and 3 hold for any function h_τ, k_τ that are continuous and bounded. In this section, we give the explicit formula for these functions that we will use in section 4, and also present our approach to find optimized solid approximating piecewise constant thermal diffusivity, also termed as multi-materials [56,43].

We first introduce the next smooth regularization of the Heaviside step function [50]:

$$\tilde{h}_\tau(y, y_0, a, b) = a + (b-a)\left(\frac{1}{1+\exp(-\tau(y-y_0))} - \frac{1}{1+\exp(\tau y_0)}\right), \tag{6}$$

where $y \in [0, y_{\max}]$. It is easy to check the following pointwise convergence

$$\lim_{\tau\to+\infty} \tilde{h}_\tau(y, y_0, a, b) = \begin{cases} a & \text{if } y < y_0, \\ (a+b)/2 & \text{if } y = y_0, \\ b & \text{if } y > y_0. \end{cases} \tag{7}$$

Hence, we set in Eq. (4)

$$h_\tau(\alpha) = \tilde{h}_\tau(\alpha, \alpha_0, 0, \alpha_{\max}). \tag{8}$$

for some $\alpha_0 \in (0, \alpha_{\max})$. Owning to (7), it has the properties wanted for an approximation of the indicator function of the solid/fluid region.

We now present the multi-material interpolation function used to search for optimized solid with piecewise constant thermal diffusivity. Our idea to approximate the thermal diffusivity constants k_j for $j = 1, \cdots, N$ is to introduce another design variable φ that interpolates the multiple values of the thermal conductivity. Therefore, k_τ is replaced with

$$k_\tau(\alpha, \varphi) = \frac{1}{k_f}\left(k_f + \tilde{h}_\tau(\alpha, \alpha_0, 0, 1)(\chi_\tau(\varphi) - k_f)\right), \tag{9}$$

where k_f is the thermal diffusivity of the fluid and χ_τ is going to interpolate the different possible diffusivities of the solid. Note that, in the fluid part of the domain, one has $\alpha(x) < \alpha_0$ and then (7) ensures that $k_\tau(\alpha, \varphi) \to 1$ as $\tau \to +\infty$ for any φ. In the solid part of the domain, $\alpha(x) \geq \alpha_0$ and (7) shows that $\tilde{h}_\tau(\alpha, \alpha_0, 0, 1) \to 1$ and thus $k_\tau(\alpha, \varphi) \to (k_f + (\chi_\tau(\varphi) - k_f))/k_f = \chi_\tau(\varphi)/k_f$ as $\tau \to +\infty$. The function $\chi_\tau(\varphi)$ is thus defined thanks to a superposition of $\tilde{h}_\tau$ as follows:

$$\chi_\tau(\varphi) = k_1 + \sum_{j=1}^{N-1} \tilde{h}_\tau(\varphi, \varphi_j, 0, a_j), \tag{10}$$

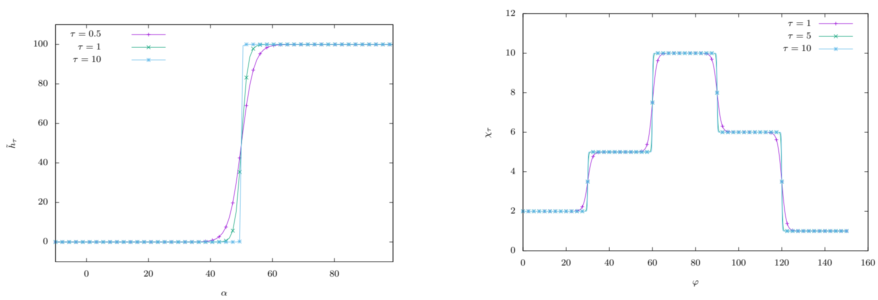
where $\varphi_j \in [0, \varphi_{\max}]$ and $\varphi_i < \varphi_j$ for any $i < j$, $i, j \in \{1, \cdots, N-1\}$. The constants a_j are then determined thanks to the following requirements

$$\lim_{\tau\to+\infty} \chi_\tau(\varphi) \ = k_j \text{ for } \varphi_{j-1} < \varphi < \varphi_j, \ 2 \leq j \leq N. \tag{11}$$

Note that (11) gives $N-1$ linear equations which determine the constants a_j. Using (7), this triangular linear system reads

$$\sum_{j=1}^{i} a_j = k_{i+1} - k_1, i = 1, 2, ..., N-1 \tag{12}$$

Figure 1 shows the behavior of the two interpolation functions h_τ and k_τ for various values of τ.



(a) $\tilde{h}_\tau(\xi, 50, 0, 100)$ for different values of τ.

(b) χ_τ for different values of τ, with $\varphi_{\max} = 150$ and interpolating values $\{2, 5, 10, 6, 1\}$.

Fig. 1: Representation of $\tilde{h}_\tau$ and χ_τ

In order to keep the analysis as general as possible, we define h_τ and k_τ as function of a parameter $\xi : \Omega \to \mathbb{R}^m$ for some integer $m \geq 1$, which will be set as $\xi = (\alpha, \varphi)$ in section 4 for the numerical applications.

Weak formulation

Before deriving a weak formulation of (4)-(5), we introduce the spaces

$$\begin{aligned} X_1^u &= \left\{ v \in H^1(\Omega)^d \mid v_{|\Gamma_\mathrm{w}} = 0 \right\}, \\ X^u &= \left\{ v \in H^1(\Omega)^d \mid v_{|(\Gamma_\mathrm{w} \cup \Gamma_\mathrm{in})} = 0 \right\}, \\ X^\theta &= \left\{ v \in H^1(\Omega) \mid v_{|\Gamma_\mathrm{in}} = 0 \right\}. \end{aligned}$$

A variational formulation of (4)-(5) then reads:

Find $(u, \theta, p) \in X_1^u \times X^\theta \times L^2(\Omega)$ such that:

$$\begin{cases} u_{|\Gamma_\mathrm{in}} = u_\mathrm{in}, & \\ a(\xi; u, v_1) + b(v_1, p) + c(u, u, v_1) + f(\theta, v_1) = 0, & \forall v_1 \in X^u \\ \tilde{a}(\xi; \theta, v_2) + \tilde{c}(\theta, u, v_2) = \displaystyle\int_{\Gamma_\mathrm{w}} C\phi v_2, & \forall v_2 \in X^\theta \\ b(u, q) = 0, & \forall q \in L^2(\Omega). \end{cases} \tag{13}$$

where

$$a(\xi; u, v_1) = \int_\Omega \left[A\nabla u : \nabla v_1 + h_\tau(\xi) u \cdot v_1\right],$$

$$b(u,q) = -\int_\Omega q(\nabla \cdot u), \quad c(u, v_1, w) = \int_\Omega (u \cdot \nabla) v_1 \cdot w, \quad f(\theta, v_1) = -\int_\Omega B\theta e_y \cdot v_1,$$

$$\tilde{a}(\xi; \theta, v_2) = \int_\Omega C k_\tau(\xi) \nabla\theta \cdot \nabla v_2, \qquad \tilde{c}(\theta, u, v_2) = \int_\Omega (\nabla\theta \cdot u) v_2.$$

Also, we endow the spaces X^u and X^θ with the following norms:

$$\|\cdot\|_{X^u}^2 = a(\xi; \cdot, \cdot), \ \|\cdot\|_{X^\theta} = \|\cdot\|_{H^1(\Omega)}.$$

Remark that, for $u \in X^u$, there exist constants $C_i(\Omega)$ depending only on Ω such that:

$$\|u\|_{L^4(\Omega)} \leq C_1(\Omega) \|u\|_{H^1(\Omega)} \leq C_2(\Omega) \|\nabla u\|_{L^2(\Omega)} \leq C_3(\Omega) \|u\|_{X^u}, \tag{14}$$

thanks to the embedding $H^1(\Omega) \subset L^4(\Omega)$ and the Poincaré inequality. Using also Hölder's inequality, we infer

$$|c(u, v, w)| \leq C_{NL} \|u\|_{X^u} \|v\|_{X^u} \|w\|_{X^u}, \ |\tilde{c}(\theta, u, v_2)| \leq C_{NL} \|u\|_{X^u} \|v_2\|_{X^\theta} \|\theta\|_{X^\theta}$$

where $C_{NL} > 0$ only depends on Ω.

Eventually, the following general TO problem is studied in this paper

$$\min \mathcal{J}(\xi, u, \theta, p)$$
$$\text{s.t.} \begin{cases} (u, \theta, p) \text{ solution of (13) parametrized by } \xi, \\ \xi \in \mathcal{U}_{ad}, \end{cases}$$

where $\mathcal{J}$ is a given cost function and $\mathcal{U}_{ad}$ is the space of vector valued bounded function with bounded variation [5, 29] on Ω, with non-negative values and a prescribed bound on its total variation, i.e. for some $\xi_{\max} \in \mathbb{R}^m$, $\xi_{\max} > 0$ and $\kappa > 0$,

$$\mathcal{U}_{ad} = \left\{\xi \in \mathrm{BV}(\Omega)^m \ : \ 0 \leq \xi(x) \leq \xi_{\max} \text{ a.e. on } \Omega, \ |D\xi|(\Omega) \leq \kappa\right\}.$$

Throughout this paper, the inequalities involving vector-valued functions are understood component-wise. There exist many physical examples that enter this framework, such as the minimization of the total pressure drop $\int_\Gamma (p + \frac{1}{2}|u|^2)(u \cdot n)$ or the maximization of the thermal exchange $\int_\Gamma \theta(u \cdot n)$, as introduced in [51].

2 Study of the PDE system and its discretization

This section will focus on the study of the underlying PDE system in the TO problem, namely on (13). We first prove the existence of a unique solution to (13) using a fixed point approach. We will afterwards analyze the finite-element discretization of (13), proving once again the existence of a unique solution and more importantly, the convergence of the discretized solution toward the continuous one.

2.1 Existence and uniqueness

We begin our analysis with some specification on our model. The next set of assumptions is supposed to hold throughout this paper.

Assumption 1 – *ξ is a vector valued bounded function of dimension $m \geq 1$ with bounded variation on Ω, has non-negative values and a prescribed bound on its total variation, i.e. for some $\xi_{max} > 0$ and $\kappa > 0$,*

$$\xi \in \mathcal{U}_{ad} = \left\{ \xi \in BV(\Omega)^m \ : \ 0 \leq \xi(x) \leq \xi_{max} \text{ a.e. on } \Omega, \ |D\xi|(\Omega) \leq \kappa \right\}.$$

Remark that $\mathcal{U}_{ad}$ is a convex, closed, weak- closed subset of $BV(\Omega)$ since the application $\alpha \in BV(\Omega) \mapsto |D\alpha|(\Omega) \in \mathbb{R}$ is lower semi-continuous (see [5, p. 120, Proposition 3.6]).*
- *We suppose that there exists $k_{min} > 0$ such that, for all $\xi \in \mathbb{R}^m$ such that $0 \leq \xi \leq \xi_{max}$, $k_{min} \leq k_\mathcal{T}(\xi)$ and $0 \leq h_\mathcal{T}(\xi)$.*
- *$h_\mathcal{T}$ and $k_\mathcal{T}$ are bounded and continuous on their domain of definition.*
- *There exists $V \in H^1(\Omega)^d$ such that $\nabla \cdot V = 0$, $V|_{\Gamma_w} = 0$, $V|_{\Gamma_{in}} = u_{in}$ and $\|V\|_{H^1(\Omega)^d} \leq M_V \|u_{in}\|_{H^{1/2}_{00}(\Gamma_{in})^d}$ for some constant $M_V > 0$.*

Concerning the last assumption, [23, Lemma 16] proves that for all $u_{\text{in}} \in H^{1/2}_{00}(\Gamma_{\text{in}})^d$, such V exists and satisfies $\|V\|_{H^1(\Omega)^d} \leq M_V \|u_{\text{in}}\|_{H^{1/2}_{00}(\Gamma_{\text{in}})^d}$ for some constant $M_V > 0$. We use this assumption to deal with the inhomogeneous Dirichlet boundary condition on Γ_{in}. One can therefore write $u = w + V$, where $w \in X^u$ satisfies:

$$\text{Find } (w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega) \text{ such that:}$$
$$\begin{cases} a(\xi; w, v_1) + b(v_1, p) = \langle G(\xi; w) + F(\theta), v_1 \rangle, & \forall v_1 \in X^u, \\ \tilde{a}(\xi; \theta, v_2) + \tilde{c}(\theta, w + V, v_2) = \left\langle \tilde{G}, v_2 \right\rangle, & \forall v_2 \in X^\theta, \\ b(w, q) = 0, & \forall q \in L^2(\Omega), \end{cases} \tag{15}$$

where:

$$\langle G(\xi; w) + F(\theta), v \rangle = -c(w + V, w + V, v) - f(\theta, v) - a(\xi, V, v),$$
$$\left\langle \tilde{G}, v \right\rangle = \int_{\Gamma_{\text{w}}} C\phi v.$$

Note that (15) is a fixed point equation equivalent to (13). To study the well-posedness of (15), we consider first the following linear problem:

$$\text{Find } (w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega) \text{ such that:}$$
$$\begin{cases} a(\xi; w, v_1) + b(v_1, p) = \langle G(\xi; \tilde{w}) + F(\theta), v_1 \rangle, & \forall v_1 \in X^u, \\ \tilde{a}(\xi; \theta, v_2) + \tilde{c}(\theta, \tilde{w} + V, v_2) = \left\langle \tilde{G}, v_2 \right\rangle, & \forall v_2 \in X^\theta, \\ b(w, q) = 0, & \forall q \in L^2(\Omega). \end{cases} \tag{16}$$

for some fixed $\tilde{w} \in X^u$. We start our analysis by proving the well-posedness of (16).

Proposition 1 *Assume that $\widetilde{w} \in X^u$ satisfies*

$$\|\tilde{w} + V\|_{X^u} \leq \frac{Ck_{min}}{1+\varepsilon},$$

for some $\varepsilon > 0$. Then problem (16) has a unique solution $(w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega)$ that satisfies

$$\|\theta\|_{X^\theta} \leq \frac{\tilde{g}_0(\Omega)(1+\varepsilon)}{\varepsilon C k_{min}} \left\|\tilde{G}\right\|_{(X^\theta)'} \leq \frac{\tilde{g}_0(\Omega)(1+\varepsilon)}{\varepsilon C k_{min}} \|\phi\|_{L^2(\Gamma)},$$

$$\|w\|_{X^u} \leq \|G(\xi;\tilde{w}) + F(\theta)\|_{(X^u)'},$$

$$\|p\|_{L^2(\Omega)} \leq \frac{2}{\beta}\|G(\xi;\tilde{w}) + F(\theta)\|_{(X^u)'},$$

where β is a positive constant.

Proof One can easily prove that the application $\tilde{b} : (\theta, v_2) \mapsto \tilde{a}(\xi; \theta, v_2) + \tilde{c}(\theta, \tilde{w} + V, v_2)$ is bilinear and satisfies the estimates:

$$|\tilde{b}(\theta, v_2)| \leq C(\Omega) \max\{\|\tilde{w} + V\|_{X^u}, C\|k_\tau\|_\infty\} \|\theta\|_{X^\theta} \|v_2\|_{X^\theta},$$

$$\tilde{b}(\theta,\theta) \geq C(\Omega) \left(Ck_{\min} - \|\tilde{w} + V\|_{X^u}\right) \|\theta\|^2_{X^\theta}.$$

Therefore, if one chooses $\tilde{w}$ such that $\|\tilde{w} + V\|_{X^u} \leq \frac{Ck_{\min}}{1+\varepsilon}$ for some $\varepsilon > 0$, one proves that $\tilde{b}$ is continuous and coercive with constant $C(\Omega)Ck_{\min}\frac{\varepsilon}{1+\varepsilon}$. Therefore, thanks to the Lax-Milgram theorem, one proves that there exists a unique function θ solving the second equation of (16) and respecting the following estimate:

$$\|\theta\|_{X^\theta} \leq \frac{1+\varepsilon}{\varepsilon C(\Omega) C k_{\min}} \left\|\tilde{G}\right\|_{(X^\theta)'} \leq \frac{\tilde{g}_0(\Omega)(1+\varepsilon)}{\varepsilon C k_{\min}} \|\phi\|_{L^2(\Gamma)},$$

where $\tilde{g}_0(\Omega)$ is a positive constant that only depends on Ω.

We are left with an equation satisfied by (w, p), which is a standard linear saddle-point problem. Since the bilinear form $a(\xi; \cdot, \cdot)$ is continuous and coercive (with respect to the norm defined on X^u), it only remains to prove that the bilinear form b satisfies an inf-sup condition. Adapting the result of [10, p.362, Eq. (2.16)], one proves that there exists a constant $\beta > 0$ such that

$$\inf_{q \in L^2(\Omega)\setminus\{0\}} \sup_{v \in X^u \setminus\{0\}} \frac{b(v,q)}{\|v\|_{X^u} \|q\|_{L^2(\Omega)}} \geq \beta.$$

We eventually conclude using [15, II.1, Proposition 1.3].

We also need some upper bound and Lipschitz condition on $G(\xi;\cdot)$ in order to state a fixed point result. This is done in the following Lemma.

Lemma 1 *The nonlinear function $G(\xi;\cdot) + F(\cdot) : X^u \times X^\theta \to (X^u)'$ satisfies the following estimates:*

$$\|G(\xi;\tilde{w}) + F(\theta)\|_{(X^u)'} \leq C \left(\|u_{in}\|^2_{H^{1/2}_{00}(\Gamma_{in})^d} + \|u_{in}\|_{H^{1/2}_{00}(\Gamma_{in})^d} + B\,\|\theta\|_{X^\theta} + \|\tilde{w}\|^2_{X^u} \right),$$

$$\|(G(\xi; w_1) + F(\theta_1)) - (G(\xi; w_2) + F(\theta_2))\|_{(X^u)'} \leq$$
$$C_L \left((\|w_1\|_{X^u} + \|w_2\|_{X^u} + \|u_{in}\|_{H^{1/2}_{00}(\Gamma_{in})^d}) \|w_1 - w_2\|_{X^u} + B\|\theta_1 - \theta_2\|_{X^\theta} \right),$$

where C, C_L are positive constants that only depend on Ω.

Proof From the inequality (14), together with the Hölder inequality, one proves: $|c(u,v,w)| \leq C_{NL}\|u\|_{X^u}\|v\|_{X^u}\|w\|_{X^u}$ where $C_{NL} > 0$ only depends on Ω. Using now the bound on V, we get

$$\begin{aligned}
|\langle G(\xi;w)+F(\theta),v\rangle_{(X^u)',X^u}| \leq & C_{NL}\|v\|_{X^u}\|w+V\|^2_{X^u} + B\|\theta\|_{X^\theta}\|v\|_{X^u}+ \\
& \|V\|_{X^u}\|v\|_{X^u} \\
\leq & \|v\|_{X^u}\Big(C_{NL}(2M_V^2\|u_{in}\|^2_{H^{1/2}_{00}(\Gamma_{\rm in})^d} + \|w\|^2_{X^u}) \\
& +B\|\theta\|_{X^\theta} + M_V\|u_{in}\|_{H^{1/2}_{00}(\Gamma_{\rm in})^d}\Big) \\
\leq & \|v\|_{X^u}C(\Omega)\Big(\|u_{in}\|^2_{H^{1/2}_{00}(\Gamma_{\rm in})^d} + \|w\|^2_{X^u}+ \\
& B\|\theta\|_{X^\theta} + \|u_{in}\|_{H^{1/2}_{00}(\Gamma_{\rm in})^d}\Big).
\end{aligned}$$

Taking the supremum over $v \in X^u$ with $\|v\|_{X^u} \leq 1$ yields the first estimates. Concerning the second estimate, note that:

$$\begin{aligned}
&\langle G(\xi;w_1)+F(\theta_1)) - (G(\xi;w_2)+F(\theta_2),v\rangle_{(X^u)',X^u} \\
&= c(w_2+V,w_2+V,v) - c(w_1+V,w_1+V,v) + f(\theta_2,v) - f(\theta_1,v) \\
&= \int_\Omega (((w_1+V)\cdot\nabla)(w_1+V) - ((w_2+V)\cdot\nabla)(w_2+V))\cdot v \\
&\quad - \int_\Omega B(\theta_2-\theta_1)e_y\cdot v
\end{aligned}$$

For two vector fields a and b, one has the following bound:

$$|a\cdot\nabla a - b\cdot\nabla b| \leq |a-b|\,|\nabla a| + |b|\,|\nabla(a-b)|\,.$$

Therefore there exists $C_L > 0$ such that:

$$\begin{aligned}
&\left|\langle (G(\xi;w_1)+F(\theta_1)) - (G(\xi;w_2)+F(\theta_2)),v\rangle_{(X^u)',X^u}\right| \\
&\leq C_L\|v\|_{X^u}\Big(\Big(\|w_1\|_{X^u} + \|w_2\|_{X^u} + \|u_{\rm in}\|_{H^{1/2}_{00}(\Gamma_{\rm in})^d}\Big)\|w_1-w_2\|_{X^u} \\
&\quad + B\|\theta_1-\theta_2\|_{X^\theta}\Big)
\end{aligned}$$

Taking once again the supremum over $v \in X^u$ with $\|v\|_{X^u} \leq 1$ finishes the proof.

Let us now move back to the non-linear problem (15). As stated before, we use the properties we have proved on (16) in order to state some fixed point result to prove the existence of solution to (15).

Theorem 1 *Given any $\xi \in \mathcal{U}_{ad}$ and given B, $\|u_{in}\|_{H^{1/2}_{00}(\Gamma_{in})}$ and $\|\phi\|_{L^2(\Gamma)}$ small enough (see* (17)-(19) *below), there exists a unique solution $(w,\theta,p) \in X^u\times X^\theta\times L^2(\Omega)$ to the variational problem* (15).

Proof Denote by $\mathcal{S}(\widetilde{w}) : (X^u)'\times(X^\theta)' \to X^u\times X^\theta\times L^2(\Omega)$, $\mathcal{S}(\widetilde{w}) : (F,G) \mapsto (w,p,\theta)$, the (linear) solution map associated to the linear problem:

Find $(w,\theta,p) \in X^u \times X^\theta \times L^2(\Omega)$ such that:

$$\begin{cases}
a(\xi;w,v_1) + b(v_1,p) = \langle F,v_1\rangle, & \forall v_1 \in X^u, \\
\tilde{a}(\xi;\theta,v_2) + \tilde{c}(\theta,\tilde{w}+V,v_2) = \langle G,v_2\rangle\,, & \forall v_2 \in X^\theta, \\
b(w,q) = 0, & \forall q \in L^2(\Omega),
\end{cases}$$

for some $(F,G) \in (X^u)' \times (X^\theta)'$. Thanks to Proposition 1, the operator $\mathcal{S}(\widetilde{w})$ is well-defined for $\|\tilde{w}+V\|_{X^u} \leq \frac{Ck_{\min}}{1+\varepsilon}$. It is also continuous and we have the estimate

$$\|\mathcal{S}(\widetilde{w})[F,G]\|_{X^u \times X^\theta} \leq C_{\mathcal{S}} \left(\|F\|_{(X^u)'} + \|G\|_{(X^\theta)'} \right),$$
$$C_{\mathcal{S}} = \max \left\{ \frac{\tilde{g}_0(\Omega)(1+\varepsilon)}{\varepsilon C k_{\min}}, 1 + \frac{2}{\beta} \right\}.$$

Problem (15) then becomes the fixed point equation:

$$(w, \theta, p) = \mathcal{T}(w, \theta, p),$$

where $\mathcal{T}(w, \theta, p) = \mathcal{S}(w)[G(\xi; w) + F(\theta), \tilde{G}]$. Denote, for $R > 0$, the set

$$B_R = \left\{ (w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega) : \|w\|_{X^u} + \|\theta\|_{X^\theta} + \|p\|_{L^2(\Omega)} \leq R \right\}.$$

Thanks to Proposition 1 and Lemma 1, we have the estimate

$$\begin{aligned} \|\mathcal{T}(w, \theta, p)\|_{X^u \times X^\theta \times L^2(\Omega)} &\leq C_{\mathcal{S}} \left(\|G(\xi; w) + F(\theta)\|_{(X^u)'} + \left\| \widetilde{G} \right\|_{(X^\theta)'} \right) \\ &\leq C_{\mathcal{S}} \left(C_1 + \|w\|_{X^u}^2 + B \|\theta\|_{X^\theta} \right), \end{aligned}$$

where

$$C_1 = C(\Omega) \left(\left\| \widetilde{G} \right\|_{(X^\theta)'} + \|u_{\text{in}}\|^2_{H^{1/2}_{00}(\Gamma_{\text{in}})^d} + \|u_{\text{in}}\|_{H^{1/2}_{00}(\Gamma_{\text{in}})^d} \right).$$

We assume $B^2 C_{\mathcal{S}}^2 - 4 C_1 C_{\mathcal{S}}^2 - 2 B C_{\mathcal{S}} + 1 \geq 0$. This reduces to

$$4 C_1 C_{\mathcal{S}}^2 \leq (B C_{\mathcal{S}} - 1)^2, \tag{17}$$

which amount to have the source terms small enough. Assuming now that $(w, \theta, p) \in B_R$ with R such that

$$R \leq R_0, \ R_0 = \frac{1 - B C_{\mathcal{S}} + \sqrt{B^2 C_{\mathcal{S}}^2 - 4 C_1 C_{\mathcal{S}}^2 - 2 B C_{\mathcal{S}} + 1}}{2 C_{\mathcal{S}}}, \tag{18}$$

we obtain that $\mathcal{T}$ maps B_R to B_R for any $R \leq R_0$.

We now prove that $\mathcal{T} : B_R \to B_R$ is a contraction mapping. Let (w_1, θ_1, p_1) and $(w_2, \theta_2, p_2) \in B_R$. We have $\mathcal{T}(w_1, \theta_1, p_1) - \mathcal{T}(w_2, \theta_2, p_2) = \mathcal{M}_1 + \mathcal{M}_2$, where

$$\mathcal{M}_1 = \{\mathcal{S}(w_1) - \mathcal{S}(w_2)\} [G(\xi; w_1) + F(\theta_1), \tilde{G}],$$
$$\mathcal{M}_2 = \mathcal{S}(w_2) \left[(G(\xi; w_1) + F(\theta_1)) - (G(\xi; w_2) + F(\theta_2)), 0 \right].$$

Lemma 1 gives the next bound for $\mathcal{M}_2$

$$\begin{aligned} \|\mathcal{M}_2\|_{X^u \times X^\theta \times L^2(\Omega)} \leq\ & C_{\mathcal{S}} C_L \max \left\{ 2R + \|u_{\text{in}}\|_{H^{1/2}_{00}(\Gamma_{\text{in}})^d}, B \right\} \\ & \times \left(\|w_1 - w_2\|_{X^u} + \|\theta_1 - \theta_2\|_{X^\theta} \right). \end{aligned}$$

The bound for $\mathcal{M}_1$ can be obtained by noting that the operator $\{\mathcal{S}(w_1) - \mathcal{S}(w_2)\} [F, G]$ verifies the identity

$$\{\mathcal{S}(w_1) - \mathcal{S}(w_2)\} [F, G] = \mathcal{S}(w_1)[0, \widetilde{C}],$$

where $\widetilde{C} \in (X^\theta)'$ is defined as $\left\langle \widetilde{C}, v_2 \right\rangle = \widetilde{c}(\theta(w_2, \theta_2, p_2), w_2 - w_1, v_2)$, with $\theta(w_2, \theta_2, p_2)$ being the temperature defined thanks to the operator $\mathcal{S}(w_2)[F, G]$. We then have the next upper bound

$$\begin{aligned}
\|\mathcal{M}_1\|_{X^u \times X^\theta \times L^2(\Omega)} &\leq C_{\mathcal{S}} \left\|\widetilde{C}\right\|_{(X^\theta)'} \leq C_{\mathcal{S}} C(\Omega) \|\theta(w_2, \theta_2, p_2)\|_{X^\theta} \|w_1 - w_2\|_{X^u} \\
&\leq C_{\mathcal{S}}^2 C(\Omega) \left(\|G(\xi, w_2) + F(\theta_2)\|_{(X^u)'} + \left\|\widetilde{G}\right\|_{(X^\theta)'} \right) \\
&\qquad \|w_1 - w_2\|_{X^u} \\
&\leq C_{\mathcal{S}}^2 C(\Omega)(C_1 + R^2 + BR) \|w_1 - w_2\|_{X^u}.
\end{aligned}$$

Gathering the previous estimates, we obtain

$$\|\mathcal{T}(w_1, \theta_1, p_1) - \mathcal{T}(w_2, \theta_2, p_2)\|_{X^u \times X^\theta \times L^2(\Omega)} \leq C_{\mathrm{Lip}} \left(\|w_1 - w_2\|_{X^u} + \|\theta_1 - \theta_2\|_{X^\theta}\right),$$

with

$$C_{\mathrm{Lip}} := \max \left\{ C_{\mathcal{S}} C_L B, C_{\mathcal{S}} C_L \left(2R + \|u_{\mathrm{in}}\|_{H_{00}^{1/2}(\Gamma_{\mathrm{in}})^d} \right), C_{\mathcal{S}}^2 C(\Omega)(C_1 + R^2 + BR) \right\}.$$

Assuming now the source terms are small enough so that

$$\|u_{\mathrm{in}}\|_{H_{00}^{1/2}(\Gamma_{\mathrm{in}})^d} < \frac{1}{C_{\mathcal{S}} C_L}, \ 4C_1 C_{\mathcal{S}}^2 C(\Omega)^2 < B^2 C_{\mathcal{S}}^2 C(\Omega)^2 + 4C(\Omega), \tag{19}$$

and that

$$\begin{aligned}
&C_{\mathcal{S}} C_L B < 1, \ R \leq \min\{R_1, R_2\}, \\
&R_1 = \frac{1}{2} \left(\frac{1}{C_{\mathcal{S}} C_L} - \|u_{\mathrm{in}}\|_{H_{00}^{1/2}(\Gamma_{\mathrm{in}})^d} \right), \\
&R_2 = \frac{1}{2C_{\mathcal{S}} C(\Omega)} \left(\sqrt{B^2 C_{\mathcal{S}}^2 C(\Omega)^2 - 4C_1 C_{\mathcal{S}}^2 C(\Omega)^2 + 4C(\Omega)} - C(\Omega) B C_{\mathcal{S}} \right).
\end{aligned} \tag{20}$$

One finally proves that for any $R \leq \min(R_0, R_1, R_2)$, $\mathcal{T} : B_R \to B_R$ is a contraction mapping if one takes B, $\|u_{\mathrm{in}}\|_{H_{00}^{1/2}(\Gamma_{\mathrm{in}})}$ and $\|\phi\|_{L^2(\Gamma)}$ small enough (see (17)-(19)). Banach fixed point theorem then proves the theorem.

Remark 1 Let us consider $h_\tau = \frac{1}{\tau} 1_{\Omega_s}$ where the solid is located in $\Omega_s \subset \Omega$. Theorem 1 ensures the existence and uniqueness of a solution to Problem (15) that satisfy the bound $\|w\|_{X^u} + \|\theta\|_{X^\theta} + \|p\|_{L^2(\Omega)} \leq R$, where R does not depend on τ (see (17)-(19)). Since $\|.\|_{X^u} = \sqrt{a(\xi; \cdot, \cdot)}$, we get $\left\|\sqrt{h_\tau} w\right\|_{L^2(\Omega)} = \frac{1}{\sqrt{\tau}} \|w\|_{L^2(\Omega_s)} \leq R$, from which we infer $\|w\|_{L^2(\Omega_s)} \leq R\sqrt{\tau}$. In addition, from the multiplicative trace inequality, we have

$$\|w\|_{L^2(\partial\Omega_s)} \leq C \sqrt{\|w\|_{L^2(\Omega_s)} \|w\|_{H^1(\Omega)}} \leq C\sqrt{R} \tau^{1/4} \sqrt{\|w\|_{X^u}} \leq C R \tau^{1/4},$$

where $C > 0$ is a generic constant. Therefore the velocity of the fluid vanishes in the solid as τ goes to 0 and it satisfies the no-slip boundary condition on $\partial\Omega_s$. It is finally worth noting that we obtain similar convergence rates as those proved in [7, Corollary 4.1, Lemma 4.4], where incompressible unsteady Navier-Stokes equations with homogeneous Dirichlet boundary on $\partial\Omega$ were considered.

2.2 Convergence of a finite element approximation

We now move on to the analysis of the discretization of (15). We consider a quasi-uniform family of triangulations (see [28, Definition 1.140]) $\{\mathcal{T}_h\}_{h>0}$ of Ω whose elements are triangles ($d=2$) or tetrahedrons ($d=3$). We emphasize that $\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} K$. The parameter h_K is the diameter of the circle or sphere inscribed in the cell $K \in \mathcal{T}_h$ and we set $h = \sup_{K\in\mathcal{T}_h} h_K$. We consider the Taylor-Hood finite element [53] that uses piecewise polynomial approximations $(w_h, \theta_h, p_h) \in X_h^u \times X_h^\theta \times M_h$ of $(w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega)$ with

$$\begin{aligned} X_h^u &= \left\{ v_h^u \in X^u \,|\, \forall K \in \mathcal{T}_h,\ v_h^u|_K \in \mathbb{P}_2(K) \right\}, \\ X_h^\theta &= \left\{ v_h^\theta \in X^\theta \,|\, \forall K \in \mathcal{T}_h,\ v_h^\theta|_K \in \mathbb{P}_2(K) \right\}, \\ M_h &= \left\{ q_h \in \mathcal{C}^0(\overline{\Omega}) \,|\, \forall K \in \mathcal{T}_h,\ q_{h\,|K} \in \mathbb{P}_1(K) \right\}, \end{aligned}$$

and we denote by $\mathcal{I}_{X_h^u} : X^u \to X_h^u$ the finite element interpolate on X_h^u. Regarding the discretization of ξ by some ξ_h, we use:

$$\mathcal{K}_h = \{\xi_h \in L^\infty(\Omega) : \xi_h|_K \in \mathbb{P}_0(T),\ \forall K \in \mathcal{T}_h\},$$

hence we consider piecewise constant polynomials over $\mathcal{T}_h$ as discrete optimization parameter.

Let us now consider the following discretized variational problem: given $\xi_h \in \mathcal{K}_h$:

$$\text{Find } (w_h, \theta_h, p_h) \in X_h^u \times X_h^\theta \times M_h \text{ such that:}$$
$$\begin{cases} a(\xi_h; w_h, v_h^u) + b(v_h^u, p_h) = \langle G(\xi_h; w_h) + F(\theta_h), v_h^u \rangle, & \forall v_h^u \in X_h^u, \\ \tilde{a}(\xi_h; \theta_h, v_h^\theta) + \tilde{c}(\theta_h, w_h + V, v_h^\theta) = \left\langle \tilde{G}, v_h^\theta \right\rangle, & \forall v_h^\theta \in X_h^\theta, \\ b(w_h, q_h) = 0, & \forall q_h \in M_h. \end{cases} \tag{21}$$

Throughout this section, we make the following assumption on $\mathcal{T}_h$:

Assumption 2 *At least an edge (d=2) or a face (d=3) of an element of $\mathcal{T}_h$ is contained in Γ_{out}.*

This assumption is fulfilled for h small enough [11]. If Assumption 2 holds, [11, Lemma 3.2], proves that there exists $\beta^* > 0$ such that:

$$\inf_{q_h \in M_h \setminus \{0\}} \sup_{v_h \in X_h^u} \frac{b(v_h, q_h)}{\|v_h\|_{X^u} \|q_h\|_{L^2(\Omega)}} \geq \beta^*.$$

Therefore, [34, Theorem 4.1], and a similar proof as for Theorem 1 (using a fixed point approach) prove that Problem (21) admits a unique solution for h small enough which satisfies:

$$\|w_h\|_{X^u} + \|\theta_h\|_{X^\theta} + \|p_h\|_{L^2(\Omega)} \leq R, \tag{22}$$

for some R that does not depend on h (see (17)-(19)).

We now prove convergence of the discretized solutions to the continuous ones. In the following results, we consider a sequence of controls ξ_h which converges to a control ξ in the weak-* topology of $\mathrm{BV}(\Omega)^m$ as $h \to 0$. This means $\xi_h \to \xi$ strongly in L^1 and $D\xi_h \xrightharpoonup{*} D\xi$ weakly-* in $\mathcal{M}_b(\Omega)$, the space of bounded Radon measure (see [5, p. 124, Definition 3.11]).

Lemma 2 *Consider a sequence of controls ξ_h which converges to a control ξ in the weak-* topology of $BV(\Omega)^m$, and suppose k_τ and h_τ to be bounded and continuous. Then there exists a subsequence $k_\tau \circ \xi_{h_k}$ (resp. $h_\tau \circ \xi_{h_k}$) which converges pointwise almost everywhere in Ω to $k_\tau \circ \xi$ (resp. $h_\tau \circ \xi$).*

Proof Since $\xi_h \to \xi$ strongly in $L^1(\Omega)^m$, there exists a subsequence of (ξ_h), denoted (ξ_{h_k}), which converges pointwise to ξ almost everywhere in Ω. Therefore, since k_τ is continuous, $k_\tau \circ \xi_{h_k}$ converges pointwise almost everywhere to $k_\tau \circ \xi$. One then proves easily that

$$k_\tau \circ \xi_{h_k} \to k_\tau \circ \xi \text{ for almost every } x \in \Omega.$$

The same proof holds true for h_τ.

Given this lemma, we prove the convergence of the finite element approximation toward the continuous solution of (15).

Theorem 2 *Suppose B, $\|u_{in}\|_{H^{1/2}_{00}(\Gamma_{in})}$ and $\|\phi\|_{L^2(\Gamma)}$ are small enough for the solutions of* (15) *and* (21) *to exist for $h > 0$ small enough, and that $\xi_h \overset{*}{\rightharpoonup} \xi$ in $BV(\Omega)^m$. Denote by (w_h, θ_h, p_h) the solution of* (21) *parametrized by ξ_h, and by (w, θ, p) the solution of* (15) *parametrized by ξ. We then have the following convergence:*

$$\lim_{h\to 0} \left(\|w_h - w\|_{X^u} + \|\theta_h - \theta\|_{X^\theta} + \|p_h - p\|_{L^2(\Omega)}\right) = 0.$$

Proof As proved in inequality (22), the sequence (w_h, θ_h, p_h) is uniformly bounded in $X^u \times X^\theta \times L^2(\Omega)$ with respect to h. Therefore, there exist $(w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega)$ and subsequences such that:

$$(w_{h_k}, \theta_{h_k}, p_{h_k}) \rightharpoonup (w, \theta, p) \text{ weakly in } H^1(\Omega)^{d+1} \times L^2(\Omega),$$

$$(w_{h_k}, \theta_{h_k}) \to (w, \theta) \text{ strongly in } L^4(\Omega)^{d+1},$$

Using also Lemma 2, we have

$$k_\tau \circ \xi_{h_k} \to k_\tau \circ \xi,\ h_\tau \circ \xi_{h_k} \to h_\tau \circ \xi \text{ for almost every } x \in \Omega.$$

Part 1 Let us prove that $(w_{h_k}, \theta_{h_k}, p_{h_k})$ weakly converges to a solution of (15) parametrized by ξ. Let $(v^u, v^\theta, q) \in \mathcal{C}^\infty(\overline{\Omega})^{d+2}$. There exists a sequence $(v^u_h, v^\theta_h, q_h) \in X^u_h \times X^\theta_h \times M_h$ that strongly converges to (v^u, v^θ, q) in $X^u \times X^\theta \times L^2(\Omega)$. Using the convergence of the subsequences, one proves: $b(v^u_{h_k}, p_{h_k}) \to b(v^u, p)$, $b(w_{h_k}, q_{h_k}) \to b(w, q)$, $c(w_{h_k} + V, w_{h_k} + V, v^u_{h_k}) \to c(w + V, w + V, v^u)$, $f(\theta_{h_k}, v^u_{h_k}) \to f(\theta, v^u)$, $\tilde{c}(\theta_{h_k}, w_{h_k} + V, v^\theta_{h_k}) \to \tilde{c}(\theta, w + V, v^\theta)$. Also, one proves:

$$\begin{aligned}
|a(\xi_{h_k}; w_{h_k}, v^u_{h_k}) - a(\xi; w, v^u)| \leq |A| & \left| \int_\Omega \nabla w_{h_k} : \nabla v_{h_k} - \nabla w : \nabla v \right| \\
& + \|h_\tau\|_\infty \left| \int_\Omega w_{h_k} \cdot v^u_{h_k} - w \cdot v^u \right| \\
& + \left| \int_\Omega (h_\tau(\xi_{h_k}) - h_\tau(\xi)) w \cdot v^u \right| \\
& \xrightarrow[k\to+\infty]{} 0,
\end{aligned}$$

$$|\tilde{a}(\xi_{h_k};\theta_{h_k},v^\theta_{h_k}) - \tilde{a}(\xi;\theta,v^\theta)| \leq C\|k_\tau\|_\infty \left|\int_\Omega \nabla\theta_{h_k}\cdot\nabla v_{h_k} - \nabla\theta\cdot\nabla v^\theta\right| + \left|\int_\Omega (k_\tau(\xi_{h_k}) - k_\tau(\xi))\nabla\theta\cdot\nabla v^\theta\right| \xrightarrow[k\to+\infty]{} 0,$$

$$\left|\int_{\Gamma_w}\phi v^\theta_{h_k} - \int_{\Gamma_\mathrm{w}}\phi v^\theta\right| \leq \left|\int_{\Gamma_\mathrm{w}}\phi(v^\theta_{h_k} - v^\theta)\right| \xrightarrow[k\to+\infty]{} 0,$$

and $a(\xi_{h_k};V,v^u_{h_k}) \xrightarrow[k\to+\infty]{} a(\xi;V,v^u)$. The previous inequalities ensure that

$$\langle G(\xi_{h_k};w_{h_k}) + F(\theta_{h_k}), v^u_{h_k}\rangle \to \langle G(\xi;w) + F(\theta), v^u\rangle.$$

It finally proves that the limit (w,θ,p) satisfies (15) for all $(v^u,v^\theta,q) \in \mathcal{C}^\infty(\overline{\Omega})^{d+2}$. The density of smooth functions in $X^u \times X^\theta \times L^2(\Omega)$ ensures that $(w,\theta,p) \in X^u \times X^\theta \times L^2(\Omega)$ satisfies (15) for all $(v^u,v^\theta,q) \in X^u \times X^\theta \times L^2(\Omega)$. Thus, $(w_{h_k},\theta_{h_k},p_{h_k})$ weakly converges toward (w,θ,p) solution of (15).

Part 2 Let us now show that w_h and θ_h strongly converge in $X^u \times X^\theta$. First, note that the application $(w,v) \mapsto a(\xi;w,v)$ defines an inner product on X^u. Taking $v^u_{h_k} = w_{h_k}$ and $v^\theta_{h_k} = \theta_{h_k}$ in (21), one gets:

$$a(\xi_{h_k};w_{h_k},w_{h_k}) = \langle G(\xi_{h_k};w_{h_k}) + F(\theta_{h_k}), w_{h_k}\rangle,$$
$$\widetilde{a}(\xi_{h_k};\theta_{h_k},\theta_{h_k}) = \langle\widetilde{G},\theta_{h_k}\rangle - \widetilde{c}(\theta_{h_k}, w_{h_k} + V, \theta_{h_k}).$$

Using now that $\xi_{h_k} \to \xi$ almost everywhere in Ω, $w_{h_k} \to w$ weakly in X^u and strongly in $L^4(\Omega)^d$, $\theta_{h_k} \to \theta$ weakly in X^θ and strongly in $L^4(\Omega)$, we obtain

$$\langle G(\xi_{h_k};w_{h_k}) + F(\theta_{h_k}), w_{h_k}\rangle \to \langle G(\xi;w) + F(\theta), w\rangle = a(\xi;w,w),$$
$$\langle\widetilde{G},\theta_{h_k}\rangle - \widetilde{c}(\theta_{h_k}, w_{h_k} + V, \theta_{h_k}) \to \langle\widetilde{G},\theta\rangle - \widetilde{c}(\theta, w + V, \theta) = \widetilde{a}(\xi;\theta,\theta),$$

where we used that $(w,\theta,p) \in X^u \times X^\theta \times L^2(\Omega)$ satisfies (15). We thus proved that

$$a(\xi_{h_k};w_{h_k},w_{h_k}) \to a(\xi;w,w),$$
$$\widetilde{a}(\xi_{h_k};\theta_{h_k},\theta_{h_k}) \to \widetilde{a}(\xi;\theta,\theta).$$

Therefore, we have proved that $w_{h_k} \rightharpoonup w$ and $a(\xi_{h_k};w_{h_k},w_{h_k}) \to a(\xi;w,w)$. This eventually proves that $w_{h_k} \xrightarrow[k\to+\infty]{} w$ strongly in X^u. Similarly, we obtain $\theta_{h_k} \xrightarrow[k\to+\infty]{} \theta$ strongly in X^θ.

Part 3 Let $\Pi_h : p \in L^2(\Omega) \mapsto \Pi_h p \in M_h$ be the finite element projector defined for all $v_h \in M_h$ as $(\Pi_h p, v_h)_{L^2(\Omega)} = (p, v_h)_{L^2(\Omega)}$. Then, we have the estimate

$$\|\Pi_h p - p\|_{L^2(\Omega)} \leq \inf_{v_h\in M_h} \|p - v_h\|_{L^2(\Omega)}$$

and the density of smooth function in $L^2(\Omega)$ ensures that $\Pi_h p$ converges toward p as $h \to 0$ strongly in $L^2(\Omega)$. Using the discrete inf-sup condition, we have

$$\|\Pi_h p - p_{h_k}\|_{L^2(\Omega)} \leq \frac{1}{\beta^*}\sup_{v_{h_k}\in B_{h_k}} \left(b(v_{h_k}, \Pi_h p - p_{h_k})\right), \tag{23}$$

where $B_h = \{u_h \in X_h^u : \|u_h\|_{X^u} = 1\}$. We emphasize that

$$\sup_{v_{h_k}\in B_{h_k}} (b(v_{h_k}, \Pi_h p - p_{h_k})) \leq \sup_{v_{h_k}\in B_{h_k}} |b(v_{h_k}, \Pi_h p - p)| + \sup_{v_{h_k}\in B_{h_k}} |b(v_{h_k}, p - p_{h_k})|,$$

and that the first term in the right hand side goes to zero as $h \to 0$. We are then left with bounding the second one. Since $(w_{h_k}, \theta_{h_k}, p_{h_k})$ satisfies (21) and (w, θ, p) satisfies (15), we can use Hölder's inequality, to get:

$$\begin{aligned}
\left|b(v_{h_k}, p) - b(v_{h_k}, p_{h_k})\right| \quad &\leq \left|a(\xi_{h_k}; w_{h_k}, v_{h_k}) - a(\xi; w, v_{h_k})\right| \\
&\quad + \left|f(\theta_{h_k}, v_{h_k}) - f(\theta, v_{h_k})\right| \\
&\quad + \left|c(w_{h_k} + V, w_{h_k} + V, v_{h_k}) - c(w + V, w + V, v_{h_k})\right| \\
&\leq \left(\int_\Omega (h_\tau(\xi_{h_k}) - h_\tau(\xi))^2\right)^{\frac{1}{2}} \|v_{h_k}\|_{X^u} \|w_{h_k}\|_{X^u} \\
&\quad + \|h_\tau\|_\infty \left|\int_\Omega (w_{h_k} - w) \cdot v_{h_k}\right| \\
&\quad + |B| \|\theta_{h_k} - \theta\|_{L^2(\Omega)} \|v_{h_k}\|_{L^2(\Omega)} \\
&\quad + \|w_{h_k} - w\|_{L^4(\Omega)} \|\nabla(w_{h_k} + V)\|_{L^2(\Omega)} \|v_{h_k}\|_{L^4(\Omega)} \\
&\quad + \|w + V\|_{L^4(\Omega)} \|\nabla(w_{h_k} - w)\|_{L^2(\Omega)} \|v_{h_k}\|_{L^4(\Omega)}.
\end{aligned}$$

Therefore, one proves that:

$$\begin{aligned}
0 \leq \sup_{v_{h_k}\in B_{h_k}} b(v_{h_k}, p - p_{h_k}) \quad &\leq \left(\int_\Omega (h_\tau(\xi_{h_k}) - h_\tau(\xi))^2\right)^{\frac{1}{2}} \|w_{h_k}\|_{X^u} \\
&\quad + \|h_\tau\|_\infty \|w_{h_k} - w\|_{L^2(\Omega)} \\
&\quad + |B| \|\theta_{h_k} - \theta\|_{L^2(\Omega)} \\
&\quad + C_{4,2} \|w_{h_k} - w\|_{X^u} \|\nabla(w_{h_k} + V)\|_{L^2(\Omega)} \\
&\quad + C_{4,2} \|w + V\|_{X^u} \|\nabla(w_{h_k} - w)\|_{L^2(\Omega)},
\end{aligned}$$

with a positive constant $C_{4,2}$. Due to the aforementioned strong convergence of (w_{h_k}, θ_{h_k}) to (w, θ) in $X^u \times X^\theta$ and lemma 2, it proves that

$$\lim_{k\to+\infty} \sup_{v_{h_k}\in B_{h_k}} (b(v_{h_k}, p) - b(v_{h_k}, p_{h_k})) = 0.$$

Eventually, using (23) and the triangular inequality, it proves that $p_{h_k} \xrightarrow[k\to+\infty]{} p$ strongly in $L^2(\Omega)$.

To summarise, we have proved that there exists a subsequence $(w_{h_k}, \theta_{h_k}, p_{h_k})$ which converges strongly to a solution (w, θ, p) of (15) when $\xi_h \overset{*}{\rightharpoonup} \xi$.

Part 4 Let us eventually prove that the whole sequence actually converges. Denote by $S_h = (w_h, \theta_h, p_h)$ a sequence of solutions and $S = (w, \theta, p)$. Since S_h is bounded, so is every subsequence S_{h_k} of S_h. Therefore, we can extract another subsequence of S_{h_k} which will also converge to S using the same arguments as in Part 1-3 and by uniqueness of the solution to (15). Therefore, every subsequence of (S_h) has a further subsequence that strongly converges to S, and using Urysohn's subsequence principle, one proves that the whole sequence $(S_h)_h$ strongly converges to S.

Remark 2 Theorem 2 gives the convergence, as $h \to 0$, of the finite element approximation of (15). There is however no additional information on the rate of convergence. Optimal error estimates can actually be obtained using results from [16] (see also [35]). Nevertheless, these require the solution to Problem (15) to be more regular (e.g. $(w, \theta, p) \in H^2(\Omega)^d \times H^2(\Omega) \times H^1(\Omega)$).

3 Optimization problem

Now that we have proved that the system (13) is well-posed, we now tackle the optimal control problem involving it, namely:

$$\begin{aligned} &\min \mathcal{J}(\xi, w, \theta, p) \\ &\text{s.t.} \begin{cases} (w, \theta, p) \text{ solution of (15) parametrized by } \xi, \\ \xi \in \mathcal{U}_{ad}, \end{cases} \end{aligned} \tag{24}$$

where $\mathcal{J}$ is a real-valued cost functional. We will follow the approach used in section 2, namely we first study the existence of an optimal solution to (24) and next its discretization. We then prove the convergence of discrete optimum toward continuous one. We end up with a proof of a necessary condition of optimality for (24).

3.1 Continuous optimization problem

We start this study with the existence of a solution to (24).

Theorem 3 *Suppose:*

(A1) $\inf_{\mathcal{U}_{ad} \times X^u \times X^\theta \times L^2(\Omega)} \mathcal{J} > -\infty$.
(A2) $\mathcal{J}$ *is lower semi-continuous w.r.t. the (weak-*, weak, weak, weak) topology of* $BV(\Omega)^m \times X^u \times X^\theta \times L^2(\Omega)$.

Then the optimization problem (24) *has at least one solution in* $\mathcal{U}_{ad} \times X^u \times X^\theta \times L^2(\Omega)$.

Proof We recall that $U_{ad} \subset \mathrm{BV}(\Omega)^m$ is a weak-* closed subset of $\mathrm{BV}(\Omega)^m$. Let $(\xi_n) \subset U_{ad}$ be a sequence uniformly bounded in U_{ad} and converging to $\xi \in BV(\Omega)^m$. One can therefore prove that $\xi_n \overset{*}{\rightharpoonup} \xi$ in U_{ad}. Let (w_n, θ_n, p_n) be the solutions of (the continuous) problem (15) parametrized by ξ_n, and (w, θ, p) the solution of (15) parametrized by ξ. Using the same technique as in the proof of theorem 2, one can get that $(w_n, \theta_n, q_n) \to (w, \theta, q)$ strongly in $X^u \times X^\theta \times L^2(\Omega)$. In other words, the mapping

$$\xi \in (U,\ \text{weak-*}) \mapsto (w, \theta, p) \in (X^u \times X^\theta \times L^2(\Omega),\ \text{strong})$$

is continuous. The proof is now based on minimizing sequence and can be adapted for instance from [36, Theorem 2.1].

3.2 Discrete optimization problem

We now turn our attention to the discretization of Problem (24), which reads:

$$\min \mathcal{J}(\xi_h, w_h, \theta_h, p_h)$$
$$\text{s.t.} \begin{cases} (w_h, \theta_h, p_h) \text{ solution of (21) parametrized by } \xi_h, \\ \xi_h \in \mathcal{U}_h = \mathcal{U}_{ad} \cap \mathcal{K}_h. \end{cases} \tag{25}$$

where $\mathcal{K}_h = \{\xi_h \in L^\infty(\Omega) : \xi_h|_T \in \mathbb{P}_0(T),\ \forall T \in \mathcal{T}_h\}$ is defined as in section 2.2. We now prove some convergence result of the finite element discretization (25) toward a solution of the continuous problem (24).

Theorem 4 *Let Assumptions (A1)-(A2) from theorem 3 be verified and assume the cost function $\mathcal{J}$ is continuous with respect to the (weak-*, strong, strong, strong) topology of $BV(\Omega)^m \times X^u \times X^\theta \times L^2(\Omega)$. Let $(\xi_h^*, w_h, \theta_h, p_h) \in \mathcal{U}_h \times X_h^u \times X_h^\theta \times M_h$ be a globally optimal solution of (25). Then $(\xi_h^*) \subset \mathcal{U}_h$ is a bounded sequence. Furthermore, there exists $(\xi^*, w^*, \theta^*, p^*) \in \mathcal{U}_{ad} \times X^u \times X^\theta \times L^2(\Omega)$ such that a subsequence of $(\xi_h^*, w_h, \theta_h, p_h)$ converges (weak-*, strong, strong, strong) to $(\xi^*, w^*, \theta^*, p^*)$ and*

$$\mathcal{J}(\xi^*, w^*, \theta^*, p^*) \le \mathcal{J}(\xi, w, \theta, p),\ \forall (\xi, w, \theta, p) \in \mathcal{U}_{ad} \times X^u \times X^\theta \times L^2(\Omega).$$

Hence, any accumulation point of $(\xi_h, w_h, \theta_h, p_h)$ is a globally optimal solution of (24).

Proof The proof can be adapted from [24, Theorem 15] (see also [37, Theorem 3]).

3.3 First order necessary conditions

Some first order necessary optimality conditions for (24) can be found in [39, Theorem 1.48]. Denote by $e(w, \theta, p, \xi) = 0$ the set of equations given by (15) and assume for now that $\partial_{(w,\theta,p)} e(w(\xi), \theta(\xi), p(\xi), \xi)$ has a bounded inverse for all $\xi \in \mathcal{U}_{\mathrm{ad}}$. Assuming that $\mathcal{J}_1 : \xi \mapsto \mathcal{J}(\xi, w(\xi), \theta(\xi), p(\xi))$ is differentiable, one gets that an optimum ξ^* satisfies:

$$\left\langle \nabla_\xi \mathcal{J}_1(\xi^*), \xi - \xi^* \right\rangle_{(L^\infty(\Omega))', L^\infty(\Omega)} \ge 0,\ \forall \xi \in U_{\mathrm{ad}}.$$

We now only need to prove that $\partial_{(w,\theta,p)} e(w(\xi), \theta(\xi), p(\xi), \xi)$ has a bounded inverse to ensure the previous inequality is valid. Thus, we study the linearization of (15). Computing the Fréchet derivative of the operator involved in (15) at $(w, \theta, p) \in X^u \times X^\theta \times L^2(\Omega)$, we have to prove that, for any source term $(f_1, f_2, f_3) \in (X^u)' \times (X^\theta)' \times L^2(\Omega)$, the linearization of (15), which reads:

Find $(\nu, \zeta, \rho) \in X^u \times X^\theta \times L^2(\Omega)$ such that

$$\begin{cases} a(\xi; \nu, v_1) + b(v_1, \rho) + c'(w + V, \nu, v_1) + f(\zeta, v_1) = \langle f_1, v_1 \rangle, & \forall v_1 \in X^u \\ \tilde{a}(\xi, \zeta, v_2) + \tilde{c}(\zeta, w + V, v_2) + \tilde{c}(\theta, \nu, v_2) = \langle f_2, v_2 \rangle & \forall v_2 \in X^\theta \\ b(\nu, q) = \langle f_3, q \rangle, & \forall q \in L^2(\Omega), \end{cases} \tag{26}$$

has a unique solution (ν, ζ, ρ) which is continuous with respect to the source terms. Above, we have

$$c'(w, \nu, v) = \int_\Omega \left((w \cdot \nabla)\nu \cdot v + (\nu \cdot \nabla) w \cdot v \right).$$

Theorem 5 *Assume u_{in} and (w,θ,p) are small enough (in norm). Then, there exists a unique solution to* (26) *which is continuous with respect to the source term (f_1,f_2,f_3).*

Proof Define the linear operators $\mathcal{A},\mathcal{B},\mathcal{F},\tilde{\mathcal{C}},\tilde{\mathcal{A}}$ as

$$\langle\mathcal{A}\nu,v_1\rangle=a(\xi;\nu,v_1)+c'(w+V,\nu,v_1),\ \langle\tilde{\mathcal{C}}\nu,v_2\rangle=\tilde{c}(\theta,\nu,v_2),$$
$$\langle\mathcal{F}\zeta,v_1\rangle=f(\zeta,v_1),\ \langle\tilde{\mathcal{A}}\zeta,v_2\rangle=\tilde{a}(\xi;\zeta,v_2)+\tilde{c}(\zeta,w+V,v_2),$$
$$\langle\mathcal{B}\rho,v_1\rangle=b(v_1,\rho).$$

Then, one can understand (26) as the following equation in $(X^u)'\times(X^\theta)'\times L^2(\Omega)$:

$$\text{Find } (\nu,\zeta,\rho)\in X^u\times X^\theta\times L^2(\Omega) \text{ such that}$$
$$\begin{cases}\mathcal{A}\nu+\mathcal{B}\rho+\mathcal{F}\zeta=f_1\\ \tilde{\mathcal{C}}\nu+\tilde{\mathcal{A}}\zeta=f_2\\ \mathcal{B}^*\nu=f_3.\end{cases}$$

Using the same analysis as in the proof of proposition 1, ζ is uniquely determined and continuous with respect to the source terms as soon as ν is. Therefore, $\zeta=\tilde{\mathcal{A}}^{-1}(f_2-\tilde{\mathcal{C}}\nu)$. We reintroduce this in the first equation, which now reads:

$$(\mathcal{A}-\mathcal{F}\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{C}})\nu+\mathcal{B}\rho=f_1-\tilde{\mathcal{A}}^{-1}f_2.$$

As we have proved in lemma 1, one gets that:

$$\begin{aligned}\langle\mathcal{A}\nu,\nu\rangle&=a(\xi;\nu,\nu)+c(w+V,\nu,\nu)+c(\nu,w+V,\nu),\\ &\geq\|\nu\|_{X^u}^2-2C_{NL}\|w+V\|_{X^u}\|\nu\|_{X^u}^2\\ &\geq(1-2C_{NL}(R+\|V\|_{X^u}))\|\nu\|_{X^u}^2,\end{aligned}$$

where R was defined in theorem 1. Also, from the proof of proposition 1, one has $\langle\tilde{\mathcal{A}}\zeta,\zeta\rangle\geq C(\Omega)(Ck_{min}-\|u+V\|_{X^u})\|\zeta\|_{X^\theta}^2\geq C(\Omega)(Ck_{min}-(R+\|V\|_{X^u}))\|\zeta\|_{X^\theta}^2$. Denote $C_{\tilde{A}}=C(\Omega)(Ck_{min}-(R+\|V\|_{X^u}))$. Since $\tilde{\mathcal{A}}$ is invertible, for all $f\in(X^\theta)'$, there exists $\zeta\in X^\theta$ such that $\tilde{\mathcal{A}}\zeta=f$ and:

$$\|\tilde{\mathcal{A}}^{-1}f\|_{X^\theta}^2\leq\frac{1}{C_{\tilde{A}}}\langle f,\tilde{\mathcal{A}}^{-1}f\rangle_{(X^\theta)',X^\theta}\leq\frac{1}{C_{\tilde{A}}}\|\tilde{\mathcal{A}}^{-1}f\|_{X^\theta}\|f\|_{(X^\theta)'}$$

which proves $\|\tilde{\mathcal{A}}^{-1}\|=\sup_{\|f\|_{(X^\theta)'}=1}\|\tilde{\mathcal{A}}^{-1}f\|_{X^\theta}\leq\frac{1}{C_{\tilde{A}}}$. Furthermore, one can easily show that: $\|\mathcal{F}\|\leq B,\ \|\tilde{\mathcal{C}}\|\leq\|\theta\|_{X^\theta}\leq R$. Eventually, all these results leads to:

$$\langle(\mathcal{A}-\mathcal{F}\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{C}})\nu,\nu\rangle\geq\left(1-2C_{NL}(R+\|V\|_{X^u})-\bar{C}\right)\|\nu\|^2,$$

where $\bar{C}=\frac{BR}{C(\Omega)(Ck_{min}-(R+\|V\|_{X^u}))}$. We recall the reader that

$$\|V\|_{X^u}\leq M_V\,\|u_{\text{in}}\|_{H_{00}^{1/2}(\Gamma_{\text{in}})}.$$

Therefore, if u_{in} has a small enough norm and if one chooses R small enough (which means that (u,θ,p) is small enough, as in theorem 1), one has that, for all $f\in(X_1^u)'$, there exists a unique $\nu\in X_1^u$ such that $\left(\mathcal{A}-\mathcal{F}\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{C}}\right)\nu=f$, thanks to Lax-Milgram theorem. Eventually, noting that (ν,ρ) satisfies a standard saddle-point problem that verifies the assumptions of [35, Theorem 4.1, p.59], one proves the existence, uniqueness and continuity with respect to the data of (ν,ζ,ρ) satisfying (26).

4 Numerical method

We focus in this section on the developments made to numerically treat problem (24). As stated in the introduction, we consider from now on that $\xi = (\alpha, \varphi)$ and h_τ and k_τ are defined as in (6) and (9). Therefore, we suppose that α and φ are scalar BV functions such that there exist positive constants $\alpha_{\max}, \varphi_{\max}$ such that $0 \leq \alpha \leq \alpha_{\max}$ and $0 \leq \varphi \leq \varphi_{\max}$ a.e. on Ω.

There exists a huge literature on the numerical methods for solving programs with PDE-constraints [8,38]. Most methods rely on the gradient of the cost functional with respect to the design variable (in our case, with respect to ξ) in order to compute a descent direction ; in turn, one needs to compute the derivative of the state variables with respect to the design variable. Since we use a *differentiate then discretize* approach, this gradient will be computed via the so-called *adjoint system* associated to (24), which will be afterward discretized.

4.1 Gradient computation with the adjoint system

From now on, we suppose that

$$\mathcal{J}(\xi, u, p, \theta) = \int_\Omega J_\Omega(\xi, u, \theta, p) + \int_\Gamma J_\Gamma(\xi, u, \theta, p),$$

where $J_\Omega : \mathcal{U}_{ad} \times X_1^u \times X^\theta \times L^2(\Omega) \to L^1(\Omega)$, $J_\Gamma : \mathcal{U}_{ad} \times X_1^u \times X^\theta \times L^2(\Omega) \to L^1(\Gamma)$ are Fréchet differentiable mappings. Slightly adapting the result of [50], one can prove that the adjoint system associated to (24) reads:

$$\begin{cases} \nabla \lambda^p - h_\tau(\xi)\lambda^u + \theta \nabla \lambda^\theta + A\Delta\lambda^u + \nabla \lambda^u \cdot u - (\lambda^u \cdot \nabla) u = -\dfrac{\partial J_\Omega}{\partial u}, \\ B\lambda^u \cdot e_y + u \cdot \nabla \lambda^\theta + \nabla \cdot \Big(C k_\tau(\xi) \nabla \lambda^\theta \Big) = -\dfrac{\partial J_\Omega}{\partial \theta}, \\ \nabla \cdot \lambda^u = -\dfrac{\partial J_\Omega}{\partial p}. \end{cases} \tag{27}$$

$$\begin{aligned} &\underline{\text{On } \Gamma_{\mathrm{w}} :} && \lambda^u \cdot t = 0,\ \lambda^u \cdot n = \frac{\partial J_\Gamma}{\partial p},\ \lambda^\theta (u \cdot n) + C k_\tau(\xi) \partial_n \lambda^\theta = \frac{\partial J_\Gamma}{\partial \theta}, \\ &\underline{\text{On } \Gamma_{\mathrm{in}} :} && \lambda^u \cdot t = 0,\ \lambda^u \cdot n = \frac{\partial J_\Gamma}{\partial p},\ \lambda^\theta = 0,\ \partial_n \lambda^p = 0, \\ &\underline{\text{On } \Gamma_{\mathrm{out}} :} && \lambda^\theta (u \cdot n) + C k_\tau(\xi) \partial_n \lambda^\theta = \frac{\partial J_\Gamma}{\partial \theta}, \\ & && \lambda^p n + \lambda^\theta \theta n + A \partial_n \lambda^u + (u \cdot n) \lambda^u = \frac{\partial J_\Gamma}{\partial u}, \end{aligned} \tag{28}$$

where t is a unit tangent vector.

The existence of adjoint solutions to (27)-(28) has been proved in theorem 5. The gradient of the cost functional then reads:

$$\begin{aligned} \nabla_\xi \mathcal{J} &= \frac{\partial J_\Omega}{\partial \xi} - \frac{\partial h_\tau}{\partial \xi}(\xi) u \cdot \lambda^u - C \frac{\partial k_\tau}{\partial \xi}(\xi) \nabla \theta \cdot \nabla \lambda^\theta && \text{on } \Omega, \\ \nabla_\xi \mathcal{J} &= \frac{\partial J_\Gamma}{\partial \xi} && \text{on } \Gamma_{\mathrm{w}}, \end{aligned} \tag{29}$$

and one has the variational optimality condition:

$$\left\langle \nabla_\xi \mathcal{J},\ \beta - \xi \right\rangle_{(L^\infty(\Omega))', L^\infty(\Omega)} \geq 0,\ \forall \beta \in \mathcal{U}_{ad}. \tag{30}$$

4.2 Numerical example

We now illustrate our method with different numerical examples, designed to test different aspects of our algorithm. The whole code is available online [1]. It has been developed using `Python` and `FEniCS/DOLFIN` [42]. The optimization procedure used here is the L-BFGS-B algorithm [18]. This algorithm approximates nicely the Hessian of the cost function, while limiting the amount of memory and computations needed to handle the iterations. Especially in this context of PDE-constrained optimization problem, it may seem important to limit the size of the data. This approach has proved to be useful [1,12,21,22,27,52].

Furthermore, in order to approximate nicely the constant thermal diffusivities with $k_\mathcal{T}$, we need to apply a threshold on α at some point of the algorithm. This is done when the algorithm fails at reducing the cost (when the steplength of the line search is too small or the cost function is too flat in the descent direction, see [18] for more details on that subject).

We test our algorithm on an ascending straight pipe heated on both sides, as sketched in figure 2. For this example, we aim at maximizing the temperature at

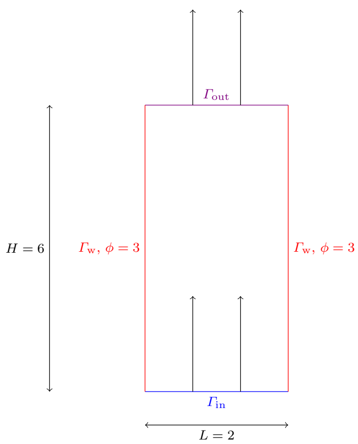


Fig. 2: Sketch of Ω

the outlet. This reads:

$$\begin{aligned} &\min -\int_{\Gamma_{\text{out}}} \theta \\ &\text{s.t.} \begin{cases} (u, \theta, p) \text{ solution of (13)}, \\ (\alpha, \varphi) \in \mathcal{U}_{ad}, \end{cases} \end{aligned}$$

[1] `https://osur-devspot.univ-reunion.fr/avieira/tossaf_pctd`

Several numerical simulations were made in order to test how the algorithm behaves when some parameters were changed, and how the cost can be influenced. The different parameters we changed were Re, Ri, τ and the size of the mesh, parameterized in `FEniCS` by the number of cells in one direction, which we denote n. One should understand that the mesh becomes thinner when n becomes larger. The default values for these parameters were

$$Re = 100,\ Ri = 1.8,\ \tau = 30,\ \alpha_{\max} = 10^8,\ \alpha_0 = 1,\ n = 40,$$

and we changed each of these parameters one by one to see the influence of it. Concerning other parameters used to define our model, the inlet velocity was defined as

$$c_{\text{in}}(x, 0) = 1.8x(2 - x),$$

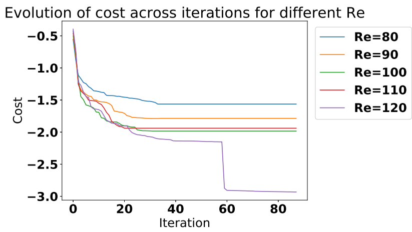
k_τ is interpolating the values $\{5, 6, 8, 11, 13, 16\}$. At the beginning of the optimization process, α is initialized at $0.15 \times x_0$ (introduced in (8)), and ϕ is initialized at ϕ_3 (introduced in (10)). Therefore, we may consider that the domain is entirely fluid at initialization.

The evolution of the cost w.r.t. the number of iterations for the different cases are shown in figs. 3-6. One can note that:

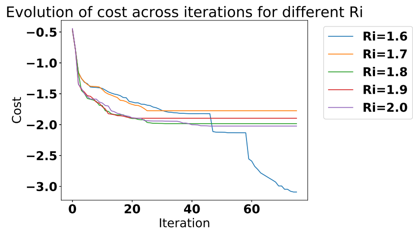
- Overall, the cost is always reduced, which shows that our algorithm works.
- Notice that the penalization model works for defining solid regions, since $u = 0$ in the zones where $h_\tau(\alpha)$ is large when comparing figure 5 and figure 6.
- As underlined in figure 3c, τ may have a significant importance in the algorithm success and must be finely tuned.
- In figure 4, we see a convergence of the minimal cost toward a value when the mesh becomes thinner, as suggested by figure 4. Note however that this theorem proves convergence of the discretized global minimum to the continuous one, and not the convergence of the minimizers. This can be seen in figure 5, which shows how the optimal k_τ changes when the mesh becomes thinner. Notice that the algorithm seems to converge towards a limit form of the solid, but the distribution of the thermal diffusivities still changes. Nonetheless, this example also shows that our approach let us optimize the distribution of the diffusivities in the solid.

5 Conclusion

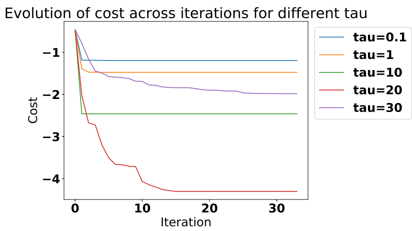
We proved the well-posedness of the penalized incompressible Navier-Stokes equations under Boussinesq assumption along with the convergence of the Taylor-Hood finite element discretization of the model. We also proved some properties of our TO problem, namely the existence of a solution and the convergence of the discretization of this problem. All these results were applied to a TO problem with materials with piecewise constant thermal diffusivity and let us design a numerical method giving interesting results, changing the design of the channel along with the thermal diffusivity. However, all these results present some limits: the source terms have to be small enough, and the penalization still contains some hard-coded parameters that need to be hand-tuned. On a final note, an other interesting study could focus on keeping the optimization problem without the penalization but rather with mixed state-control constraints that could be non-smooth, but more robust to the a priori chosen parameters.



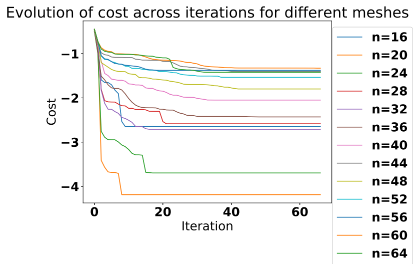
(a) Re.



(b) Ri.



(c) τ.



(d) n.

Fig. 3: Evolution of the cost w.r.t. the iterations for several values of different parameters.

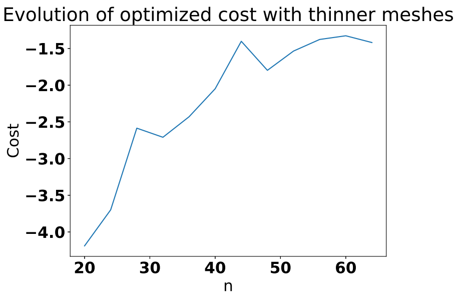


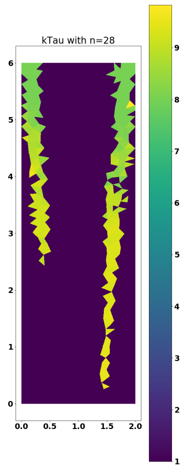
Fig. 4: Evolution of the final optimized cost w.r.t. the size of the mesh (the bigger n, the thinner the mesh).

Declaration

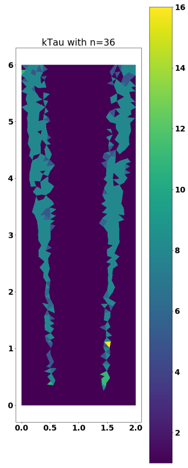
All the authors are supported by the "Agence Nationale de la Recherche" (ANR), Project O-TO-TT-FU number ANR-19-CE40-0011. The used code for this article is available at `https://osur-devspot.univ-reunion.fr/avieira/tossaf_pctd`.

References

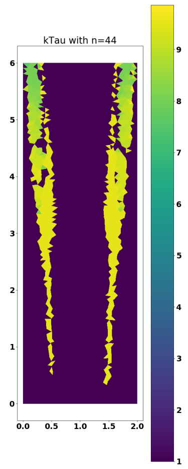
1. Akçelik, V., Biros, G., Ghattas, O., Hill, J., Keyes, D., van Bloemen Waanders, B.: Parallel algorithms for PDE-constrained optimization. In: Parallel processing for scientific computing, pp. 291–322. SIAM (2006)



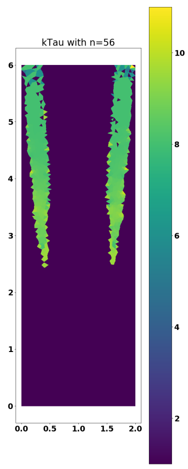
(a) $n = 28$.



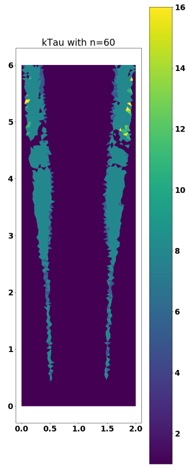
(b) $n = 36$.



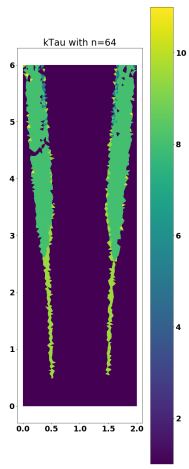
(c) $n = 44$.



(d) $n = 56$.

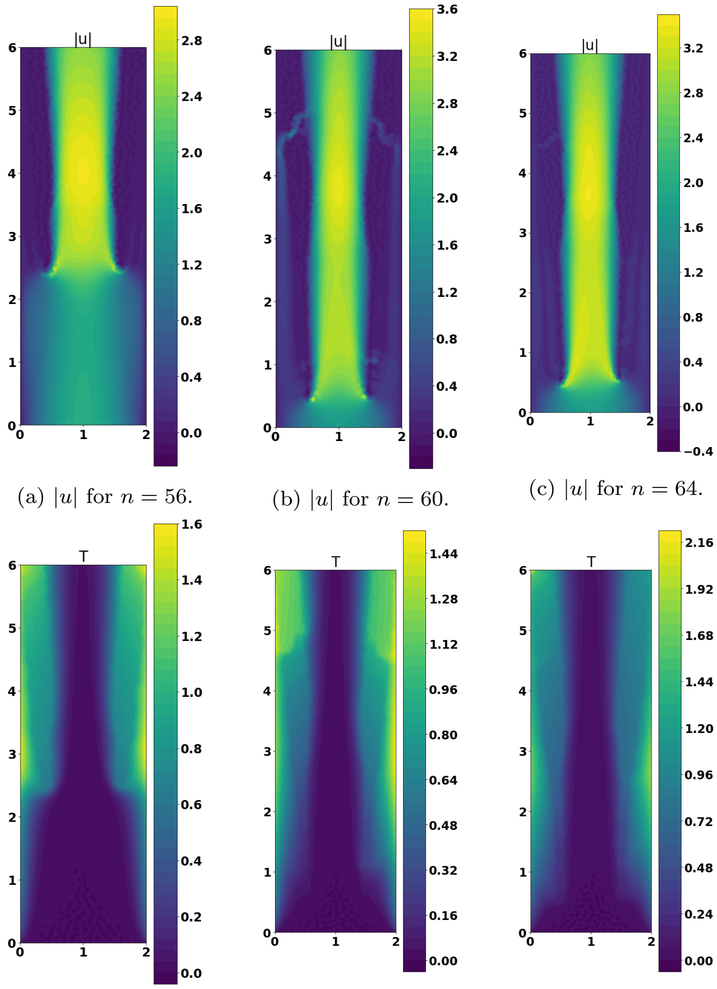


(e) $n = 60$.



(f) $n = 64$.

Fig. 5: Evolution of the optimal k_τ for several values of n.



(a) $|u|$ for $n = 56$. (b) $|u|$ for $n = 60$. (c) $|u|$ for $n = 64$.

(d) T for $n = 56$. (e) T for $n = 60$. (f) T for $n = 64$.

Fig. 6: Evolution of the optimal u and T for several values of n.

2. Alexandersen, J., Aage, N., Andreasen, C.S., Sigmund, O.: Topology optimisation for natural convection problems. International Journal for Numerical Methods in Fluids **76**(10), 699–721 (2014)
3. Alexandersen, J., Andreasen, C.S.: A review of topology optimisation for fluid-based problems. Fluids **5**(1), 29 (2020)
4. Alexandersen, J., Sigmund, O., Aage, N.: Large scale three-dimensional topology optimisation of heat sinks cooled by natural convection. International Journal of Heat and Mass Transfer **100**, 876–891 (2016)
5. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems, vol. 254. Clarendon Press Oxford (2000)
6. Amstutz, S.: The topological asymptotic for the Navier-Stokes equations. ESAIM: Control, Optimisation and Calculus of Variations **11**(3), 401–425 (2005)
7. Angot, P., Bruneau, C.H., Fabrie, P.: A penalization method to take into account obstacles in incompressible viscous flows. Numerische Mathematik **81**(4), 497–520 (1999)

8. Antil, H., Kouri, D.P., Lacasse, M.D., Ridzal, D.: Frontiers in PDE-constrained Optimization, vol. 163. Springer (2018)
9. Banks, T.H., Kunisch, K.: Estimation techniques for distributed parameter systems. Springer Science & Business Media (2012)
10. Bernardi, C., Canuto, C., Maday, Y.: Spectral approximations of the Stokes equations with boundary conditions on the pressure. SIAM Journal on Numerical Analysis **28**(2), 333–362 (1991)
11. Bernardi, C., Chacón Rebollo, T., Yakoubi, D.: Finite element discretization of the Stokes and Navier–Stokes equations with boundary conditions on the pressure. SIAM Journal on Numerical Analysis **53**(3), 1256–1279 (2015)
12. Biros, G., Ghattas, O.: Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. part i: The Krylov–Schur solver. SIAM Journal on Scientific Computing **27**(2), 687–713 (2005)
13. Boland, J., Layton, W.: Error analysis for finite element methods for steady natural convection problems. Numerical Functional Analysis and Optimization **11**(5-6), 449–483 (1990)
14. Borrvall, T., Petersson, J.: Topology optimization of fluids in Stokes flow. International Journal for Numerical Methods in Fluids **41**(1), 77–107 (2003)
15. Brezzi, F., Fortin, M.: Mixed and hybrid finite element methods. New York: Springer-Verlag (1991)
16. Brezzi, F., Rappaz, J., Raviart, P.A.: Finite dimensional approximation of nonlinear problems. Numerische Mathematik **36**(1), 1–25 (1980)
17. Bruns, T.E.: Topology optimization of convection-dominated, steady-state heat transfer problems. International Journal of Heat and Mass Transfer **50**(15-16), 2859–2873 (2007)
18. Byrd, R.H., Lu, P., Nocedal, J., Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM Journal on scientific computing **16**(5), 1190–1208 (1995)
19. Caubet, F., Conca Rosende, C., Godoy, M.: On the detection of several obstacles in 2D Stokes flow: topological sensitivity and combination with shape derivatives. IPI : Inverse Problems and Imaging (2016)
20. Chen, Z., Zou, J.: An augmented lagrangian method for identifying discontinuous parameters in elliptic systems. SIAM Journal on Control and Optimization **37**(3), 892–910 (1999)
21. Cioaca, A., Alexe, M., Sandu, A.: Second-order adjoints for solving PDE-constrained optimization problems. Optimization methods and software **27**(4-5), 625–653 (2012)
22. Cipolla, S., Durastante, F.: Fractional PDE constrained optimization: An optimize-then-discretize approach with L-BFGS and approximate inverse preconditioning. Applied Numerical Mathematics **123**, 43–57 (2018)
23. Cocquet, P.H., Rakotobe, M., Ramalingom, D., Bastide, A.: Error analysis for the finite element approximation of the darcy-brinkman-forchheimer model for porous media with mixed boundary conditions. Journal of Computational and Applied Mathematics p. 113008 (2020)
24. Cocquet, P.H., Riffo, S., Salomon, J.: Optimization of bathymetry for long waves with small amplitude. Accepted for publication in SICON (2020)
25. Colmenares, E., Gatica, G.N., Oyarzúa, R.: A posteriori error analysis of an augmented fully-mixed formulation for the stationary Boussinesq model. Computers & Mathematics with Applications **77**(3), 693–714 (2019)
26. Dbouk, T.: A review about the engineering design of optimal heat transfer systems using topology optimization. Applied Thermal Engineering **112**, 841–854 (2017)
27. Durastante, F., Cipolla, S.: Fractional PDE constrained optimization: Box and sparse constrained problems. In: Numerical Methods for Optimal Control Problems, pp. 111–135. Springer (2018)
28. Ern, A., Guermond, J.L.: Theory and practice of finite elements, vol. 159. Springer Science & Business Media (2013)
29. Evans, L.C., Gariepy, R.F.: Measure theory and fine properties of functions. Chapman and Hall/CRC (2015)
30. Evgrafov, A.: The limits of porous materials in the topology optimization of Stokes flows. Applied Mathematics and Optimization **52**(3), 263–277 (2005)
31. Feppon, F., Allaire, G., Bordeu, F., Cortial, J., Dapogny, C.: Shape optimization of a coupled thermal fluid–structure problem in a level set mesh evolution framework. SeMA Journal **76**(3), 413–458 (2019)

32. Galdi, G.: An introduction to the mathematical theory of the Navier-Stokes equations: Steady-state problems. Springer Science & Business Media (2011)
33. Garcke, H., Hinze, M., Kahle, C., Lam, K.F.: A phase field approach to shape optimization in Navier–Stokes flow with integral state constraints. Advances in Computational Mathematics **44**(5), 1345–1383 (2018)
34. Girault, V., Raviart, P.A.: Finite element approximation of the Navier-Stokes equations. Lecture Notes in Mathematics **749** (1979)
35. Girault, V., Raviart, P.A.: Finite element methods for Navier-Stokes equations: theory and algorithms, vol. 5. Springer Science & Business Media (2012)
36. Gunzburger, M.D., Hou, S.L.: Finite-dimensional approximation of a class of constrained nonlinear optimal control problems. SIAM Journal on Control and Optimization **34**(3), 1001–1043 (1996)
37. Haslinger, J., Mäkinen, R.A.: On a topology optimization problem governed by two-dimensional Helmholtz equation. Computational Optimization and Applications **62**(2), 517–544 (2015)
38. Herzog, R., Kunisch, K.: Algorithms for PDE-constrained optimization. GAMM-Mitteilungen **33**(2), 163–176 (2010)
39. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints, vol. 23. Springer Science & Business Media (2008)
40. Hinze, M., Quyen, T.N.T.: Finite element approximation of source term identification with TV-regularization. Inverse Problems **35**(12), 124004 (2019)
41. Hvejsel, C.F., Lund, E.: Material interpolation schemes for unified topology and multi-material optimization. Structural and Multidisciplinary Optimization **43**(6), 811–825 (2011)
42. Logg, A., Wells, G.N.: DOLFIN: Automated Finite Element Computing. ACM Transactions on Mathematical Software **37**(2) (2010)
43. Long, K., Wang, X., Gu, X.: Multi-material topology optimization for the transient heat conduction problem using a sequential quadratic programming algorithm. Engineering optimization **50**(12), 2091–2107 (2018)
44. Mohammadi, B., Pironneau, O.: Shape optimization in fluid mechanics. Annual Review of Fluid Mechanics **36**, 255–279 (2004)
45. Mohammadi, B., Pironneau, O.: Applied shape optimization for fluids. Oxford University Press (2010)
46. Novotny, A.A., Sokołowski, J., Żochowski, A.: Topological derivatives of shape functionals. part II: first-order method and applications. Journal of Optimization Theory and Applications **180**(3), 683–710 (2019)
47. Otárola, E., Quyen, T.N.T.: A reaction coefficient identification problem for fractional diffusion. Inverse Problems **35**(4), 045010 (2019)
48. Othmer, C.: A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows. International Journal for Numerical Methods in Fluids **58**(8), 861–877 (2008)
49. Othmer, C., Grahs, T.: Approaches to fluid dynamic optimization in the car development process. In: Proceedings of the EUROGEN Conference. Munich (2005)
50. Ramalingom, D., Cocquet, P.H., Bastide, A.: A new interpolation technique to deal with fluid-porous media interfaces for topology optimization of heat transfer. Computers & Fluids **168**, 144–158 (2018)
51. Ramalingom, D., Cocquet, P.H., Maleck, R., Bastide, A.: A multi-objective optimization problem in mixed and natural convection for a vertical channel asymmetrically heated. Structural and Multidisciplinary Optimization **60**(5), 2001–2020 (2019)
52. Schulz, V.H., Siebenborn, M., Welker, K.: Efficient PDE constrained shape optimization based on steklov–poincaré-type metrics. SIAM Journal on Optimization **26**(4), 2800–2819 (2016)
53. Taylor, C., Hood, P.: A numerical solution of the Navier-Stokes equations using the finite element technique. Computers & Fluids **1**(1), 73–100 (1973)
54. Temam, R.: Navier-Stokes equations: theory and numerical analysis, vol. 343. American Mathematical Soc. (2001)
55. Wein, F., Chen, N., Iqbal, N., Stingl, M., Avila, M.: Topology optimization of unsaturated flows in multi-material porous media: application to a simple diaper model. Communications in Nonlinear Science and Numerical Simulation **78**, 104871 (2019)
56. Zuo, W., Saitou, K.: Multi-material topology optimization using ordered SIMP interpolation. Structural and Multidisciplinary Optimization **55**(2), 477–491 (2017)