Deutsche Forschungsgemeinschaft
Priority Programme 1962

A Phase Field Approach to Shape Optimization in Navier–Stokes Flow with Integral State Constraints

Harald Garcke, Michael Hinze, Christian Kahle, Kei Fong Lam

SPP 1962
Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-059

received on June 19, 2018

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

Adv Comput Math
https://doi.org/10.1007/s10444-018-9586-8

A phase field approach to shape optimization in Navier–Stokes flow with integral state constraints

Harald Garcke[1] · **Michael Hinze**[2] · **Christian Kahle**[3] · **Kei Fong Lam**[4]

Received: 14 February 2017 / Accepted: 3 January 2018

Abstract We consider the shape optimization of an object in Navier–Stokes flow by employing a combined phase field and porous medium approach, along with additional perimeter regularization. By considering integral control and state constraints, we extend the results of earlier works concerning the existence of optimal shapes and the derivation of first order optimality conditions. The control variable is a phase field function that prescribes the shape and topology of the object, while the state variables are the velocity and the pressure of the fluid. In our analysis, we cover a multitude of constraints which include constraints on the center of mass, the volume of the fluid region, and the total potential power of the object. Finally, we present numerical results of the optimization problem that is solved using the variable metric projection type (VMPT) method proposed by Blank and Rupprecht, where we

Communicated by: Jon Wilkening

✉ Harald Garcke
Harald.Garcke@mathematik.uni-regensburg.de

Michael Hinze
Michael.Hinze@uni-hamburg.de

Christian Kahle
Christian.Kahle@ma.tum.de

Kei Fong Lam
kflam@math.cuhk.edu.hk

[1] Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany

[2] Fachbereich Mathematik, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

[3] Zentrum Mathematik, Technische Universität München, Garching bei München, Germany

[4] Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Published online: 29 January 2018

consider one example of topology optimization without constraints and one example of maximizing the lift of the object with a state constraint, as well as a comparison with earlier results for the drag minimization.

Keywords Topology optimization · Shape optimization · Phase field approach · Navier–Stokes flow · Integral state constraints

Mathematics Subject Classification (2010) 35Q35 · 35Q56 · 35R35 · 49Q10 · 49Q12 · 65M22 · 65M60 · 76S05

1 Introduction

Fundamental to the design of aircraft and cars, as well as any technologies that would involve an object traveling within a fluid, such as wind turbines and drug delivery in biomedical applications, is the consideration of hydrodynamic forces acting on the object, for example the drag and lift forces. The desire to construct an object with minimal drag or with maximal lift-to-drag ratio naturally leads to the notion of shape optimization in fluids, in which the problem can often be formulated in terms of an optimal control problem with PDE constraints.

Let us assume that $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded domain with Lipschitz boundary, and contains a non-permeable object B. We will denote the boundary of B by $\Gamma := \partial B \cap \Omega$ with the outer unit normal ν, and assume that $\Gamma \cap \partial\Omega = \emptyset$, i.e., the object B never touches the external boundary . A fluid is present in the complement region $E := \Omega \setminus B$, and we assume that the velocity $\boldsymbol{u}$ and the pressure p of the fluid in the region E obey the stationary Navier–Stokes equations with no-slip conditions on Γ, namely,

$$-\mu\Delta\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \ \text{ in } E, \tag{1a}$$

$$\operatorname{div}\boldsymbol{u} = 0 \ \text{ in } E, \tag{1b}$$

$$\boldsymbol{u} = \boldsymbol{0} \ \text{ on } \Gamma, \tag{1c}$$

$$\boldsymbol{u} = g \ \text{ on } \partial E \cap \partial\Omega. \tag{1d}$$

Here $\boldsymbol{f}$ denotes the external body force, μ denotes the (constant) viscosity, and g models the inflow and outflow on the boundary $\partial\Omega$ such that $\int_{\partial\Omega} g \cdot \nu_{\partial\Omega}\, \mathrm{d}\mathcal{H}^{d-1} = 0$, where $\nu_{\partial\Omega}$ denotes the outer unit normal on $\partial\Omega$. Our present contribution is motivated from a previous numerical study [14] for the shape optimization problem of maximizing the lift-to-drag ratio subject to the PDE constraint (1). In two spatial dimensions, the classical formulation of the lift-to-drag ratio is defined as

$$\frac{\int_\Gamma \boldsymbol{u}_\infty^\perp \cdot \left(\mu\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top\right) - p\,\mathrm{I}\right)\nu\,\mathrm{d}\mathcal{H}^{d-1}}{\int_\Gamma \boldsymbol{u}_\infty \cdot \left(\mu\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top\right) - p\,\mathrm{I}\right)\nu\,\mathrm{d}\mathcal{H}^{d-1}}, \tag{2}$$

where $\boldsymbol{u}_\infty$ is the flow direction, $\boldsymbol{u}_\infty^\perp$ is the perpendicular vector and $\mathcal{H}^{d-1}$ is the Hausdorff measure on the set Γ. In [14], using a phase field approximation which we will detail below, the authors obtain an optimal shape similar to a non-symmetric airfoil with a small angle of attack. However, a chief obstacle to a rigorous mathematical

treatment of the problem is that it is unknown if the lift-to-drag ratio (2) is bounded from above (as we want to maximize the ratio). Furthermore, due to the fractional form of the lift-to-drag ratio (2), we also observe fractions entering in the associated adjoint system and optimality conditions computed by the formal Lagrangian method, leading to severe complications in the numerical implementation.

One idea is to study a related problem involving maximizing the lift while the drag is constrained to be below a certain threshold, i.e.,

$$\begin{aligned} &\max \int_\Gamma \boldsymbol{u}_\infty^\perp \cdot \left(\mu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top\right) - p\,\mathrm{I}\right) \nu \,\mathrm{d}\mathcal{H}^{d-1} \\ &\text{subject to (1) and } \int_\Gamma \boldsymbol{u}_\infty \cdot \left(\mu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top\right) - p\,\mathrm{I}\right) \nu \,\mathrm{d}\mathcal{H}^{d-1} \leq D, \end{aligned}$$

where $D > 0$ is a threshold for the drag. In this case, the problematic fractional form is replaced and analysis can be performed on the optimization problem. In exchange we now have to deal with (integral) state constraints, and the difficulty lies in establishing the existence of the associated Lagrange multipliers.

To fix the setting for this paper, we now introduce a design function $\varphi : \Omega \to \{\pm 1\}$, where $\{\varphi = 1\} = E$ describes the fluid region and $\{\varphi = -1\} = B$ is its complement. The natural function space for the design functions is the space of bounded variations that take values ± 1, i.e., $\varphi \in BV(\Omega, \{\pm 1\})$, which implies that the fluid region E has finite perimeter $P_\Omega(E)$. If φ is a function of bounded variation, its distributional derivative $\mathrm{D}\varphi$ is a finite Radon measure which can be decomposed into a positive measure $|\,\mathrm{D}\varphi\,|$ and a S^{d-1}-valued function $\nu_\varphi \in L^1(\Omega, |\,\mathrm{D}\varphi\,|)^d$, where S^{d-1} denotes the $(d-1)$-dimensional sphere. The total variation for $\varphi \in BV(\Omega, \{\pm 1\})$, denoted by $|\,\mathrm{D}\varphi\,|(\Omega)$, satisfies

$$|\,\mathrm{D}\varphi\,|(\Omega) = 2P_\Omega(\{\varphi = 1\}),$$

and thus we can express the Hausdorff measure $\mathcal{H}^{d-1}$ on the set Γ as $\frac{1}{2}|\,\mathrm{D}\varphi\,|(\Omega)$. Furthermore, the S^{d-1}-valued function ν_φ can be considered as a generalized normal on the set $\partial\{\varphi = 1\}$ (see [1, 10, 16] for a more detailed introduction to the theory of sets of finite perimeter and functions of bounded variation).

For functions $b : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $h : \Omega \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, we consider the following general shape optimization problem with perimeter regularization:

$$\begin{aligned} \min_{(\varphi,u,p)} \mathcal{J}_0(\varphi, \boldsymbol{u}, p) \;:=\; &\textstyle\int_\Omega b(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi)\,\mathrm{dx} \\ &+ \textstyle\int_\Omega \frac{1}{2} h(x, \nabla \boldsymbol{u}, p, \nu_\varphi)\,\mathrm{d}\,|\,\mathrm{D}\varphi\,| + \frac{\gamma}{2}|\,\mathrm{D}\varphi\,|(\Omega), \end{aligned} \tag{3}$$

subject to $\varphi \in BV(\Omega, \{\pm 1\})$ and $(\boldsymbol{u}, p) \in H^1(E) \times L^2(E)$ fulfilling

$$-\mu \Delta \boldsymbol{u} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \;\text{ in } E = \{\varphi = 1\}, \tag{4a}$$

$$\operatorname{div} \boldsymbol{u} = 0 \;\text{ in } E, \tag{4b}$$

$$\boldsymbol{u} = g \;\text{ on } \partial\Omega \cap \partial E, \tag{4c}$$

$$\boldsymbol{u} = \boldsymbol{0} \;\text{ on } \Gamma = \Omega \cap \partial E. \tag{4d}$$

In addition, for fixed $m_1, m_2 \in \mathbb{N} \cup \{0\}$, we impose the m_1 integral equality constraints and m_2 integral inequality constraints:

$$G_i(\varphi, \boldsymbol{u}, p) = 0 \text{ for } 1 \le i \le m_1, \;\; G_i(\varphi, \boldsymbol{u}, p) \ge 0 \text{ for } m_1 + 1 \le i \le m_1 + m_2, \tag{5}$$

where for each $1 \le i \le m_1 + m_2$,

$$\begin{aligned} G_i(\varphi, \boldsymbol{u}, p) := & \int_\Omega K_i(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi) \,\mathrm{d}x \\ & + \int_\Gamma \tfrac{1}{2} \boldsymbol{L}_i(x, \nabla \boldsymbol{u}, p) \cdot \nu_\varphi \,\mathrm{d}\,|\,\mathrm{D}\varphi\,|\,, \end{aligned} \tag{6}$$

for functions $K_i : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{L}_i : \Omega \times \mathbb{R}^{d\times d} \times \mathbb{R} \to \mathbb{R}^d$. The parameter $\gamma > 0$ in (3) is the weighting factor for the perimeter regularization, which is given by the term $\frac{1}{2}\,|\,\mathrm{D}\varphi\,|$, and in light of the above discussion regarding the measure $\frac{1}{2}\,|\,\mathrm{D}\varphi\,|$ representing the Hausdorff measure on Γ, we see that the functions b and $\{K_i\}_{i=1}^{m_1+m_2}$ model objectives and constraints in the bulk phases E and B, while h and $\{\boldsymbol{L}_i \cdot \nu_\varphi\}_{i=1}^{m_1+m_2}$ model constraints on the interface Γ. Examples of b, h, K_i and $\boldsymbol{L}_i$ are given below, and it is noteworthy to point out that there is no dependence on $\boldsymbol{u}$ in $\boldsymbol{L}_i$ as the no-slip condition (4d) ensures that $\boldsymbol{u} = \boldsymbol{0}$ on Γ. However, the gradient $\nabla \boldsymbol{u}$ may not vanish on Γ, which leads to its appearance in the surface constraints.

The appearance of the perimeter regularization $\frac{\gamma}{2}\,|\,\mathrm{D}\varphi\,|\,(\Omega)$ in (3) is motivated from the well-known difficulties regarding the mathematical treatment of shape optimization - in particular the existence of minimizers/optimal shapes are not guaranteed [23, 26, 36]. However, if the shape optimization problem is additionally supplemented with a perimeter regularization, then positive results concerning existence of optimal shapes have been obtained (see for instance [34]).

Let us now give some examples of functions b, h, K and $\boldsymbol{L}$ (where we neglect the index i for convenience) that are of relevance. For a subset $A \subset \Omega$, we use the notation $\chi_A(x)$ to denote the characteristic function of A, i.e., $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \in \Omega \setminus A$. In particular, one can think of the design function φ as $\varphi(x) = -1 + 2\chi_E(x)$ which satisfies $\varphi(x) = 1$ for $x \in E$ and $\varphi(x) = -1$ for $x \in \Omega \setminus E = B$. Hence, in the following examples for the function b, we can use $\frac{1}{2}(1+\varphi)$ as a restriction to the region E and similarly, $\frac{1}{2}(1-\varphi)$ as a restriction to the region B:

– the total potential power of the fluid

$$\frac{1+\varphi}{2}\left(\frac{\mu}{2}\,|\,\nabla \boldsymbol{u}\,|^2 - \boldsymbol{f} \cdot \boldsymbol{u}\right), \tag{7}$$

– the construction cost of the object $\frac{1-\varphi}{2} w(x)$, where w denotes a cost function per unit volume,
– the least square approximation

$$\frac{1+\varphi}{2} \chi_{\mathcal{Q}}(x)(\delta_1\,|\,p - p_{\text{tar}}\,|^2 + \delta_2\,|\,\boldsymbol{u} - \boldsymbol{u}_{\text{tar}}\,|^2)$$

to a target velocity profile $\boldsymbol{u}_{\text{tar}}$ and a target pressure profile p_{tar} in an observation region $\mathcal{Q} \subset E$. Here δ_1 and δ_2 denote non-negative constants.

An example for the surface cost h which has practical applications is the hydrodynamic force component in the direction of the unit vector a, which is given as

$$a \cdot \left(\mu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top \right) - p\, \mathrm{I} \right) \nu_\varphi, \tag{8}$$

where I denotes the identity tensor. The drag of the object is given when a is parallel to the flow direction $\boldsymbol{u}_\infty$, while the lift of the object is given when $a = \boldsymbol{u}_\infty^\perp$, the unit vector perpendicular to the flow direction.

Examples of integral constraints that are of interests include

- volume constraints on the amount of fluid - setting $K_1 = \varphi - \beta_1$, $\boldsymbol{L}_1 = \boldsymbol{0}$, $K_2 = \beta_2 - \varphi$ and $\boldsymbol{L}_2 = \boldsymbol{0}$ for fixed constants $-1 < \beta_1 \leq \beta_2 < 1$ leads to inequality constraints:

$$G_1(\varphi) = \int_\Omega \varphi - \beta_1 \,\mathrm{dx} \geq 0, \quad G_2(\varphi) = \int_\Omega \beta_2 - \varphi \,\mathrm{dx} \geq 0,$$

  or equivalently

$$\begin{array}{l} \frac{\beta_1+1}{2} \,|\, \Omega \,| \leq \int_\Omega \frac{1+\varphi}{2} \,\mathrm{dx} = |\, E \,| \leq \frac{\beta_2+1}{2} \,|\, \Omega \,| \\ \Leftrightarrow \quad \beta_1 \,|\, \Omega \,| \leq \int_\Omega \varphi \,\mathrm{dx} \leq \beta_2 \,|\, \Omega \,|, \end{array} \tag{9}$$

- the prescribed mass of the object - setting $K_1 = M \,|\, \Omega \,|^{-1} - \frac{1-\varphi}{2}\rho(x)$, $\boldsymbol{L}_1 = \boldsymbol{0}$, where $\rho(x)$ is a mass density and $M > 0$ is a target/maximal mass leads to the inequality constraint

$$G_1(\varphi) = M - \int_\Omega \tfrac{1}{2}\rho(x)(1-\varphi) \,\mathrm{dx} \geq 0 \Leftrightarrow \int_\Omega \tfrac{1}{2}\rho(x)(1-\varphi) \,\mathrm{dx} \leq M, \tag{10}$$

- the prescribed center of mass of the object (with uniform mass density) at a point y in the interior of Ω, i.e., $y \notin \partial\Omega$ - setting $K_i = \frac{1-\varphi}{2}(x_i - y_i)$ and $\boldsymbol{L}_i = \boldsymbol{0}$ for $1 \leq i \leq d$ leads to the equality constraints

$$G_i(\varphi) = \int_\Omega \tfrac{1}{2}(1-\varphi)(x_i - y_i) \,\mathrm{dx} = 0 \text{ for } i = 1, 2, \ldots, d, \tag{11}$$

- the prescribed drag of the object - setting $\boldsymbol{L}_1 = -\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top)a + pa$, $K_1 = D \,|\, \Omega \,|^{-1}$, where a is the unit vector parallel to the flow direction $\boldsymbol{u}_\infty$ and $D > 0$ is a maximal drag value leads to the inequality constraint

$$\begin{array}{l} G_1(\varphi, \boldsymbol{u}, p) = D - a \cdot \int_\Omega \left(\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top)\nu_\varphi - p\nu_\varphi \right) \frac{1}{2} \,\mathrm{d}\, |\, \mathrm{D}\varphi \,| \geq 0 \\ \Leftrightarrow \quad a \cdot \int_\Omega \left(\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top)\nu_\varphi - p\nu_\varphi \right) \frac{1}{2} \,\mathrm{d}\, |\, \mathrm{D}\varphi \,| \leq D. \end{array} \tag{12}$$

In the examples of the cost functional described above, the problem involving minimizing the drag of the object has received much attention and is well-studied in the literature, see [3, 6, 28, 29, 32] and the references therein. For the formal derivation of shape derivatives with general volume and boundary objective functionals in Navier–Stokes flow, we refer the reader to [31], but to the authors' best knowledge, the shape optimization problem with integral state constraint has not received much attention.

In this work, we study a phase field approximation of the problem (3)-(6), which is derived in Section 2. Under appropriate assumptions we prove the existence

of minimizers to the phase field optimization problem and derive the first order optimality conditions. The main difficulty we encounter is establishing the existence of Lagrange multipliers, which is achieved via constraint qualifications such as the Zowe–Kurcyusz constraint qualification [39], for some of the integral state constraints mentioned above. We give two examples: one involves constraints that depend only on the design function φ which are the volume (9), the prescribed mass (10) and the center of mass (11), while the second example involves the total potential power (7). We encounter some technical difficulties regarding the drag constraint (12), and can only show the existence of Lagrange multipliers if the threshold D is sufficiently large. Via numerical simulations we give a proof of concept showing that with the help of the phase field approach shape and topology optimization for fluid flow taking state constraints can be solved. For large Reynolds number problems more efficient numerical solution methods have to be devised in the future.

The rest of the paper is organized as follows: In Section 2, we present the phase field approximation of (3)-(6) that utilizes the porous-medium approach of Borrvall and Petersson [7], and state several preliminary results on the state equations. Then, in Section 3, we state the assumptions on b, h, K_i and $\boldsymbol{L}_i$ that allows us to establish the existence of minimizers to the phase field shape optimization problem. Sufficient conditions on the differentiability of b, h, K_i and $\boldsymbol{L}_i$ are outlined in Section 4 which lead to the existence of Lagrange multipliers, the solvability of the adjoint system, and the derivation of the necessary optimality conditions. We verify the aforementioned conditions in Section 5 for two specific examples of integral constraints; the first example involves constraints on the mass, center of mass and volume, while the second example involves a constraint on the total potential power. Lastly, in Section 6 we briefly outline our numerical approach to solving the optimality conditions, and present several numerical simulations.

2 Phase field formulation

One approach to tackle shape optimization problems that can yield rigorous mathematical results is to employ a phase field approximation, similar in spirit to Bourdin and Chambolle [8] that was applied to topology optimization (see also [4, 27, 35, 38] and the reference cited therein), and has been recently used for drag minimization in stationary Stokes flow [12] and in stationary Navier–Stokes flow [11, 13, 14, 24].

The approach we take in this paper is similar to the previous works [11, 12, 14], in which we relax the condition that the design function φ takes only values in $\{\pm 1\}$ (i.e., $\varphi \in BV(\Omega, \{\pm 1\})$) and now allow φ to be a function with values in $\mathbb{R}$ and inherits $H^1(\Omega)$ regularity. In particular, we change the admissible space of design functions from subsets of $BV(\Omega, \{\pm 1\})$ to subsets of $H^1(\Omega)$. This leads to the development of interfacial layers $\{-1 < \varphi < 1\}$ in between the fluid region $E = \{\varphi = 1\}$ and the object region $B = \{\varphi = -1\}$. This interfacial layer replaces the boundary Γ of B and a parameter $\varepsilon > 0$ is associated to the thickness of the interfacial layer. The idea is to reformulate the original shape optimization problem (3)-(6) by taking into account

the above modification of the design functions. For the perimeter regularization, we can use the scaled Ginzburg–Landau energy functional

$$\frac{1}{2c_0}\mathcal{E}_\varepsilon(\varphi) = \frac{1}{2c_0}\int_\Omega \frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2 + \frac{1}{\varepsilon}\Psi(\varphi)\,\mathrm{dx}\,,$$

where Ψ is a potential with equal minima at $\varphi = \pm 1$, to approximate the perimeter functional P_Ω. The positive constant c_0 is dependent only on the potential Ψ via the relation

$$c_0 := \frac{1}{2}\int_{-1}^{1}\sqrt{2\Psi(s)}\,\mathrm{ds}\,, \tag{13}$$

and it is well-known that $\frac{1}{2c_0}\mathcal{E}_\varepsilon$ approximates $\varphi \mapsto \frac{1}{2}\,|\,\mathrm{D}\varphi\,|\,(\Omega) = P_\Omega(\{\varphi = 1\})$ in the sense of Γ-convergence [25].

By introducing an interfacial region between the fluid and the object, we have relaxed the non-permeability assumption of the object in the vicinity of its boundary. Therefore, we use the so-called porous medium approach of Borrvall and Petersson [7] and replace the object B with a porous medium of small permeability $(\overline{\alpha_\varepsilon})^{-1} \ll 1$. A function $\alpha_\varepsilon(\varphi)$ is introduced to interpolate between the inverse permeabilities of the fluid region $\alpha_\varepsilon(1) = 0$ and the porous medium $\alpha_\varepsilon(-1) = \overline{\alpha_\varepsilon}$, which satisfies

$$\overline{\alpha_\varepsilon} \to \infty \text{ as } \varepsilon \to 0.$$

With this, we extend the state equations from E to the whole domain Ω by the addition of the *porous-medium* term $\alpha_\varepsilon(\varphi)\boldsymbol{u}$:

$$\alpha_\varepsilon(\varphi)\boldsymbol{u} - \mu\Delta\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \ \text{ in } \Omega, \tag{14a}$$

$$\operatorname{div}\boldsymbol{u} = 0 \ \text{ in } \Omega, \tag{14b}$$

$$\boldsymbol{u} = g \ \text{ on } \partial\Omega. \tag{14c}$$

We note that this additional term vanishes in the fluid region, and in the limit $\varepsilon \to 0$, one expects the velocity $\boldsymbol{u}$ in the object region to vanish. We point out that in the modified state (14) we solve for a velocity field $\boldsymbol{u}$ and pressure field p that are defined on the fixed domain Ω. Furthermore, in the objective functional (3) and in the integral state constraints (6), the terms

$$\int_\Omega b(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p, \varphi)\,\mathrm{dx} \ \text{ and } \int_\Omega K_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p, \varphi)\,\mathrm{dx}$$

require no modification when we consider the phase field setting. For the surface terms such as

$$\int_\Omega \frac{1}{2}h(x, \nabla\boldsymbol{u}, p, \nu_\varphi)\,\mathrm{d}\,|\,\mathrm{D}\varphi\,| \ \text{ and } \int_\Omega \frac{1}{2}\boldsymbol{L}_i(x, \nabla\boldsymbol{u}, p)\cdot\nu_\varphi\,\mathrm{d}\,|\,\mathrm{D}\varphi\,|$$

arising from the objective functional (3) and the integral state constraints (6), we employ the idea in [14] to reformulate them. Assuming the function h is one-homogeneous with respect to its last variable, which is true for the case of the hydrodynamic force (8), we use the vector-valued measure with density $\frac{1}{2}\nabla\varphi$ as an

approximation to $\frac{1}{2}\nu_\varphi \, \mathrm{d} \, |\, \mathrm{D}\varphi \,|$, see [14, §3.2] for more details. This in turn gives us the phase field approximation

$$\int_\Omega \frac{1}{2} h(x, \nabla \boldsymbol{u}, p, \nabla \varphi) \, \mathrm{dx}$$

to the surface objective involving the function h and from this point onward we will assume that h is one-homogeneous with respect to its last variable. By a similar argument, we see that

$$\int_\Omega \frac{1}{2} \boldsymbol{L}_i(x, \nabla \boldsymbol{u}, p) \cdot \nabla \varphi \, \mathrm{dx}$$

is a phase field approximation of the surface integral constraint involving $\boldsymbol{L}_i$.

Recall that in the modified state (14) the porous-medium term $\alpha_\varepsilon(\varphi)\boldsymbol{u}$ serves to enforce the condition that the velocity $\boldsymbol{u}$ in the object region should vanish in the limit $\varepsilon \to 0$. In earlier works motivated by the paper of Borrvall and Petersson [7] the authors of [12, 18] also added to the phase field objective functional

$$\int_\Omega b(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi) + \frac{1}{2} h(x, \nabla \boldsymbol{u}, p, \nabla \varphi) + \frac{\gamma}{2c_0} \left(\frac{\varepsilon}{2} |\, \nabla \varphi \,|^2 + \frac{1}{\varepsilon} \Psi(\varphi) \right) \mathrm{dx}$$

a penalization term

$$\int_\Omega \frac{1}{2} \hat{\alpha}_\varepsilon(\varphi) \, |\, \boldsymbol{u} \,|^2 \, \mathrm{dx}\,, \tag{15}$$

where $\hat{\alpha}_\varepsilon$ is a function with similar properties as α_ε, i.e., $\hat{\alpha}_\varepsilon(1) = 0$ and $\hat{\alpha}_\varepsilon(-1) \to \infty$ as $\varepsilon \to 0$. In fact, in the rigorous analysis of the phase field approximation in Stokes flow, the addition of penalization term (15) to the objective functional does indeed lead to the velocity field vanishing in the object region as $\varepsilon \to 0$ (see [12, §3] and [18, §6.3] for more details). In this paper we consider including both elements in the analytical treatment of the optimization problem. It is also possible to consider $\hat{\alpha}_\varepsilon = 0$, however in this case no rigorous results on the sharp interface limit $\varepsilon \to 0$ are known.

Before we state the phase field optimization problem, let us mention that for the analysis we assume that the function α_ε in the porous-medium term satisfies the properties:

(A0) $\alpha_\varepsilon \in C^{1,1}(\mathbb{R})$ is non-negative, and there exist constants $s_a, s_b \in \mathbb{R}$ with $s_a \leq -1$ and $s_b \geq 1$ such that

$$\begin{aligned} \alpha_\varepsilon(s) &= \alpha_\varepsilon(s_a) \quad \forall s \leq s_a, \\ \alpha_\varepsilon(s) &= \alpha_\varepsilon(s_b) \quad \forall s \geq s_b. \end{aligned} \tag{16}$$

In particular, for arbitrary ϕ and its truncation $\tilde{\phi} := \max(s_a, \min(s_b, \phi))$ we see that $\alpha_\varepsilon(\phi) = \alpha_\varepsilon(\tilde{\phi})$, and so the state (14) for ϕ and $\tilde{\phi}$ are equivalent. Hence, without loss of generality, we now search for optimal design functions φ exhibiting $H^1(\Omega)$-regularity and satisfies the pointwise bounds $s_a \leq \varphi \leq s_b$ a.e. in Ω.

Taking into account the above discussions, we arrive at the following phase field approximation to the optimal control problem (3)-(6):

$$\begin{aligned}\min_{(\varphi,\boldsymbol{u},p)} \mathcal{J}_\varepsilon(\varphi,\boldsymbol{u},p) := &\int_\Omega \tfrac{1}{2}\hat{\alpha}_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2 + b(x,\boldsymbol{u},\nabla\boldsymbol{u},p,\varphi)\,\mathrm{dx}\\ &+\int_\Omega \tfrac{1}{2}h(x,\nabla\boldsymbol{u},p,\nabla\varphi) + \tfrac{\gamma}{2c_0}\left(\tfrac{1}{\varepsilon}\Psi(\varphi) + \tfrac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\mathrm{dx}\,,\end{aligned} \tag{17}$$

subject to

$$\begin{aligned}&\varphi \in \Phi := \{f \in H^1(\Omega)\,|\,s_a \le f \le s_b \text{ a.e. in } \Omega\} \subset H^1(\Omega)\cap L^\infty(\Omega),\\ &\boldsymbol{u} \in \boldsymbol{H}^1_{g,\sigma}(\Omega) := \left\{h \in \boldsymbol{H}^1(\Omega)\,|\,\operatorname{div} h = 0 \text{ in } \Omega \text{ and } h = g \text{ on } \partial\Omega\right\},\\ &p \in L^2_0(\Omega) := \left\{h \in L^2(\Omega)\,\Big|\,\int_\Omega h\,\mathrm{dx} = 0\right\}\end{aligned}$$

satisfying the weak formulation of (14a):

$$\int_\Omega \alpha_\varepsilon(\varphi)\boldsymbol{u}\cdot v + \mu\nabla\boldsymbol{u}\cdot\nabla v + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot v - p\operatorname{div} v - \boldsymbol{f}\cdot v\,\mathrm{dx} \tag{18}$$

for all $v \in H^1_0(\Omega) := \{h \in H^1(\Omega)\,|\,h = \boldsymbol{0} \text{ on } \partial\Omega\}$, along with m_1 equality and m_2 inequality integral constraints

$$G_j(\varphi,\boldsymbol{u},p) = 0 \text{ for } 1 \le j \le m_1, \quad G_{m_1+k}(\varphi,\boldsymbol{u},p) \ge 0 \text{ for } 1 \le k \le m_2,$$

of the form

$$G_i(\varphi,\boldsymbol{u},p) = \int_\Omega K_i(x,\boldsymbol{u},\nabla\boldsymbol{u},p,\varphi) + \tfrac{1}{2}\nabla\varphi\cdot\boldsymbol{L}_i(x,\nabla\boldsymbol{u},p)\,\mathrm{dx} \tag{19}$$

for $1 \le i \le m_1 + m_2$.

2.1 Preliminaries on the state equations

Since the porous medium Navier–Stokes (14) have been analyzed in detail in previous works [11, 18], we recall some useful results in this section.

Lemma 1 ([14, Lem. 4.3]) *Suppose $\alpha_\varepsilon \in C^{1,1}(\mathbb{R})$ is non-negative and satisfies (16), for every $\varphi \in L^1(\Omega)$ there exists at least one pair $(\boldsymbol{u},p) \in \boldsymbol{H}^1_{g,\sigma}(\Omega)\times L^2_0(\Omega)$ such that (18) is satisfied. Furthermore, there exists a positive constant $C = C(\mu,\alpha_\varepsilon,\boldsymbol{f},g,\Omega)$ independent of φ such that*

$$\|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega)} + \|p\|_{L^2(\Omega)} \le C. \tag{20}$$

The estimate (20) can be established by testing the weak form (18) with $\boldsymbol{u} - G$, where $G \in H^1_{g,\sigma}(\Omega)$ is a vector field depending on the inflow/outflow g and the domain Ω, and satisfying certain properties. Furthermore, this computation also shows that the constant C depends on ε only through the function α_ε.

By the above existence result, we can define a set-valued solution operator

$$S_\varepsilon(\varphi) := \left\{(\boldsymbol{u},p) \in \boldsymbol{H}^1_{g,\sigma}(\Omega)\times L^2_0(\Omega)\,|\,(\boldsymbol{u},p) \text{ satisfies (18)}\right\} \tag{21}$$

for any $\varphi \in L^1(\Omega)$. Next, we state a continuity property of the solution operator.

Lemma 2 ([14, Lem. 4.4 and 4.5]) *Let $(\varphi_k)_{k\in\mathbb{N}} \subset L^1(\Omega)$ be a sequence with corresponding solution $(\boldsymbol{u}_k, p_k) \in S_\varepsilon(\varphi_k) \subset \boldsymbol{H}^1(\Omega) \times L^2(\Omega)$ for each $k \in \mathbb{N}$. Suppose there exists $\varphi \in L^1(\Omega)$ such that*

$$\|\varphi_k - \varphi\|_{L^1(\Omega)} \to 0 \text{ as } k \to \infty.$$

Then, there exists a subsequence, denoted by the same index, and functions $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, $p \in L^2(\Omega)$ such that

$$\|\boldsymbol{u}_k - \boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega)} \to 0, \quad \|p_k - p\|_{L^2(\Omega)} \to 0 \text{ as } k \to \infty,$$

with the property that $(\boldsymbol{u}, p) \in S_\varepsilon(\varphi)$. Furthermore, it holds that

$$\lim_{k\to\infty} \int_\Omega \alpha_\varepsilon(\varphi_k) \,|\, \boldsymbol{u}_k \,|^2 \,\mathrm{dx} = \int_\Omega \alpha_\varepsilon(\varphi) \,|\, \boldsymbol{u} \,|^2 \,\mathrm{dx}\,.$$

In general, we do not have uniqueness of solutions to (18), but there is a conditional uniqueness result.

Lemma 3 ([11, Lem. 5], [18, Lem. 12.2]) *If there exists $\boldsymbol{u} \in S_\varepsilon(\varphi)$ with*

$$\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} < \frac{\mu}{K_\Omega}, \tag{22}$$

where

$$K_\Omega := \begin{cases} \frac{1}{2} \,|\, \Omega \,|^{\frac{1}{2}} & \text{for } d = 2, \\ \frac{2\sqrt{2}}{3} \,|\, \Omega \,|^{\frac{1}{6}} & \text{for } d = 3. \end{cases} \tag{23}$$

Then, $S_\varepsilon(\varphi) = \{(\boldsymbol{u}, p)\}$. That is, there is exactly one solution of (18) corresponding to $\varphi \in L^1(\Omega)$.

The additional assumption (22) on the solution $\boldsymbol{u} \in S_\varepsilon(\varphi)$ to ensure uniqueness of the state equations can be achieved for small data $\boldsymbol{f}$ and g or with high viscosity μ. However, there are also settings in which (22) can be justified a posteriorly [11]. For the subsequent analysis, more precisely in showing the differentiability of the solution operator S_ε and the derivation of the optimality conditions, we require that S_ε is a one-to-one mapping. Hence, throughout the rest of the paper we assume that (22) holds. Alternatively, instead of assuming (22), we can work with an isolated local solution to (18), for which the subsequent analysis is valid in a neighborhood of this isolated local solution.

We now state the Fréchet differentiability of the solution operator S_ε.

Lemma 4 ([14, Lem. 4.8]) *Let $\varphi_\varepsilon \in H^1(\Omega) \cap L^\infty(\Omega)$ be given such that $S_\varepsilon(\varphi_\varepsilon) = \{(\boldsymbol{u}_\varepsilon, p_\varepsilon)\}$. Then, there exists a neighborhood N of φ_ε in $H^1(\Omega) \cap L^\infty(\Omega)$ such that for every $\delta \in N$, the solution operator consists of exactly one pair, and we may write $S_\varepsilon : N \to \boldsymbol{H}^1(\Omega) \times L^2(\Omega)$. This mapping is differentiable at φ_ε with*

$$\mathrm{D}S_\varepsilon(\varphi_\varepsilon)(\delta) =: (\boldsymbol{w}_\varepsilon, r_\varepsilon) \in \boldsymbol{H}^1_0(\Omega) \times L^2_0(\Omega),$$

where $(\boldsymbol{w}_\varepsilon, r_\varepsilon)$ *is the unique solution to*

$$\begin{aligned} &\int_\Omega \alpha_\varepsilon'(\varphi_\varepsilon)\delta\boldsymbol{u}_\varepsilon \cdot v + \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{w}_\varepsilon \cdot v + \mu\nabla\boldsymbol{w}_\varepsilon \cdot \nabla v \,\mathrm{dx} \\ &\quad + \int_\Omega (\boldsymbol{w}_\varepsilon \cdot \nabla)\boldsymbol{u}_\varepsilon \cdot v + (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{w}_\varepsilon \cdot v - r_\varepsilon \operatorname{div} v \,\mathrm{dx} = 0 \quad \forall v \in H_0^1(\Omega), \end{aligned} \tag{24}$$

which is the weak formulation of the following the linearized state system

$$\begin{aligned} \alpha_\varepsilon'(\varphi_\varepsilon)\delta\boldsymbol{u}_\varepsilon + \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{w}_\varepsilon - \mu\Delta\boldsymbol{w}_\varepsilon + (\boldsymbol{w}_\varepsilon \cdot \nabla)\boldsymbol{u}_\varepsilon + (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{w}_\varepsilon + \nabla r_\varepsilon &= \boldsymbol{0} \ \textit{in } \Omega, \\ \operatorname{div} \boldsymbol{w}_\varepsilon &= 0 \ \textit{in } \Omega, \\ \boldsymbol{w}_\varepsilon &= \boldsymbol{0} \ \textit{on } \partial\Omega. \end{aligned}$$

3 Existence of a minimizer

We make the following assumptions for the potential Ψ and the functions $\hat{\alpha}_\varepsilon$, b, h, K_i, and $\boldsymbol{L}_i$.

(A1) Let $\Psi \in C^{1,1}(\mathbb{R})$ be a non-negative function such that $\Psi(s) = 0$ if and only if $s = \pm 1$, and that there exist positive constants c_1, c_2, t_0 such that

$$c_1 t^k \le \Psi(t) \le c_2 t^k \quad \forall\, |\, t \,| \ge t_0, k \ge 2.$$

(A2) The function $\hat{\alpha} \in C^{1,1}(\mathbb{R})$ satisfies the same assumptions as α_ε, i.e., $\hat{\alpha}_\varepsilon$ is non-negative, with $\hat{\alpha}_\varepsilon(1) = 0$, $\hat{\alpha}_\varepsilon(-1) \to \infty$ as $\varepsilon \to 0$, and fulfills (16).

(A3) The function $b : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function of the form

$$b(x, \boldsymbol{w}, \boldsymbol{A}, s, t) := B(x, \boldsymbol{w}, \boldsymbol{A}, s)\, z(x, t), \tag{25}$$

for some Carathéodory functions $B : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \to \mathbb{R}$, $z : \Omega \times \mathbb{R} \to \mathbb{R}$ and there exist non-negative functions $f_b \in L^1(\Omega)$, $\{f_{b,i}\}_{i=1}^4 \subset L^\infty(\Omega)$ such that for a.e. $x \in \Omega$ it holds for any $r \ge 0$, $p \ge 2$ in two-dimensions and $2 \le p \le 6$ in three-dimensions,

$$\begin{aligned} |\, B(x, \boldsymbol{w}, \boldsymbol{A}, s)\,| &\le f_b(x) + f_{b,1}(x)\,|\,\boldsymbol{w}\,|^p + f_{b,2}(x)\,|\,\boldsymbol{A}\,|^2 + f_{b,3}(x)\,|\,s\,|^2, \\ |\, z(x,t)\,| &\le f_{b,4}(x)\,|\,t\,|^r, \end{aligned}$$

for all $s, t \in \mathbb{R}$, $\boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{A} \in \mathbb{R}^{d\times d}$.

(A4) The function $h : \Omega \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a Carathéodory function that is one-homogeneous with respect to its last variable and there exist non-negative functions $f_h \in L^1(\Omega)$, $\{f_{h,k}\}_{k=1}^3 \subset L^\infty(\Omega)$ such that for a.e. $x \in \Omega$ it holds,

$$|\, h(x, \boldsymbol{A}, s, \boldsymbol{w})\,| \le f_h(x) + f_{h,1}(x)\,|\,\boldsymbol{A}\,|^2 + f_{h,2}(x)\,|\,s\,|^2 + f_{h,3}(x)\,|\,\boldsymbol{w}\,|^2,$$

for all $s, t \in \mathbb{R}$, $\boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{A} \in \mathbb{R}^{d\times d}$.

(A5) For each $1 \le i \le m_1$, the function $K_i : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function of the form

$$K_i(x, \boldsymbol{w}, \boldsymbol{A}, s, t) := \mathcal{K}_i(x, \boldsymbol{w}, \boldsymbol{A}, s)\; y_i(x, t) + k_i(x), \tag{26}$$

for functions $k_i \in L^1(\Omega)$ and Carathéodory functions $\mathcal{K}_i : \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R} \to \mathbb{R}$, $y_i : \Omega \times \mathbb{R} \to \mathbb{R}$ and there exist non-negative functions $z_1 \in L^1(\Omega)$,

$z_2, z_3, z_4, z_5 \in L^\infty(\Omega)$ such that for a.e. $x \in \Omega$ it holds for any $r \geq 0$, $p \in [2, \infty)$ in two-dimensions and $p \in [2, 6)$ in three-dimensions,

$$|\mathcal{K}_i(x, \boldsymbol{w}, \boldsymbol{A}, s)| \leq z_1(x) + z_2(x)\,|\boldsymbol{w}|^p + z_3(x)\,|\boldsymbol{A}|^2 + z_4(x)\,|s|^2,$$
$$|y_i(x, t)| \leq z_5(x)\,|t|^r,$$

for all $s, t \in \mathbb{R}$, $\boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{A} \in \mathbb{R}^{d\times d}$.

(A6) For each $1 \leq j \leq m_2$, the function $\boldsymbol{L}_j : \Omega \times \mathbb{R}^{d\times d} \times \mathbb{R} \to \mathbb{R}^d$ is a Carathéodory function and there exist non-negative functions $z_6 \in L^1(\Omega)$, $z_7, z_8 \in L^\infty(\Omega)$ such that for a.e. $x \in \Omega$ it holds,

$$\left|\boldsymbol{L}_j(x, \boldsymbol{A}, s)\right| \leq z_6(x) + z_7(x)\,|\boldsymbol{A}| + z_8(x)\,|s|,$$

for all $s \in \mathbb{R}$ and $\boldsymbol{A} \in \mathbb{R}^{d\times d}$.

(A7) We assume that the set

$$\mathbb{K}_{ad} := \Big\{\varphi \in \Phi \ \text{ with } (\boldsymbol{u}, p) = S_\varepsilon(\varphi) \text{ s.t. } G_i(\varphi, \boldsymbol{u}, p) = 0 \text{ for } 1 \leq i \leq m_1,$$
$$\text{and } G_{m_1+j}(\varphi, \boldsymbol{u}, p) \geq 0 \ \text{ for } 1 \leq j \leq m_2\Big\},$$

is non-empty, where the phase field integral state constraints G_k, for $1 \leq k \leq m_1 + m_2$, are of the form (19) and we recall the set Φ is defined as $\{f \in H^1(\Omega) \,|\, s_a \leq f \leq s_b \text{ a.e. in } \Omega\}$.

(A8) The functionals $\mathcal{B} : H^1(\Omega) \times \boldsymbol{H}^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ and $\mathcal{H} : H^1(\Omega) \times \boldsymbol{H}^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ defined as

$$\mathcal{B}(\varphi, \boldsymbol{u}, p) := \int_\Omega B(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p)\, z(x, \varphi)\,\mathrm{d}x,$$
$$\mathcal{H}(\varphi, \boldsymbol{u}, p) := \int_\Omega \frac{1}{2} h(x, \nabla \boldsymbol{u}, p, \nabla \varphi)\,\mathrm{d}x$$

satisfy $\mathcal{B}|_{\mathbb{K}_{ad}\times \boldsymbol{H}^1_{g,\sigma}(\Omega)\times L^2_0(\Omega)}$ and $\mathcal{H}|_{\mathbb{K}_{ad}\times \boldsymbol{H}^1_{g,\sigma}(\Omega)\times L^2_0(\Omega)}$ are bounded from below, $\mathcal{B}$ is weakly lower semicontinuous, and for all $\varphi_n \rightharpoonup \varphi$ in $H^1(\Omega)$, $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\boldsymbol{H}^1(\Omega)$, $p_n \to p$ in $L^2(\Omega)$, it holds that

$$\mathcal{H}(\varphi, \boldsymbol{u}, p) \leq \liminf_{n\to\infty} \mathcal{H}(\varphi_n, \boldsymbol{u}_n, p_n).$$

The particular forms of b and K_i are motivated from the discussions in Section 1, where z and y would typically be functions of the form $\frac{1+\varphi}{2}$, and the function k_i would be of the form $D\,|\,\Omega\,|^{-1}$. Furthermore, the set $\mathbb{K}_{ad}$ is the set of admissible design functions whose elements satisfy the m_1 equality integral constraints and m_2 inequality integral constraints. While we assume the non-emptiness of $\mathbb{K}_{ad}$ for the general setting here, later in Section 5 we show for two examples that the corresponding set $\mathbb{K}_{ad}$ is indeed non-empty.

The following weakly closed property is useful for showing the existence of minimizers to the optimal control problem (17)-(19).

Lemma 5 *Under* (A5) *and* (A6)*, let* $\{\varphi_n\}_{n\in\mathbb{N}}$ *be a sequence in* $\mathbb{K}_{ad}$ *such that* $\varphi_n \rightharpoonup \varphi \in H^1(\Omega)$ *for some* $\varphi \in H^1(\Omega)$*, then* $\varphi \in \mathbb{K}_{ad}$.

Proof Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{K}_{ad}$ with weak limit $\varphi \in H^1(\Omega)$. It suffices to show that if $(\boldsymbol{u}, p) \in S_\varepsilon(\varphi)$, then $G_i(\varphi, \boldsymbol{u}, p) = 0$ for $1 \le i \le m_1$ and $G_{m_1+j}(\varphi, \boldsymbol{u}, p) \ge 0$ for $1 \le j \le m_2$, which then implies that $\varphi \in \mathbb{K}_{ad}$.

Let $\{(\boldsymbol{u}_n, p_n)\}_{n\in\mathbb{N}} \subset H^1_{g,\sigma}(\Omega) \times L^2_0(\Omega)$ be the corresponding solutions to (14) for φ_n, i.e., for each $n \in \mathbb{N}$, $(\boldsymbol{u}_n, p_n) \in S_\varepsilon(\varphi_n)$. Since $\varphi_n \rightharpoonup \varphi \in H^1(\Omega)$, by compactness we have strong convergence along subsequences $\varphi_{n_j} \to \varphi$ in $L^p(\Omega)$ for $p \in [1, \infty)$ in two dimensions and $p \in [1, 6)$ in three dimensions. Consequently, we also have $\varphi_{n_j} \to \varphi$ a.e. in Ω and hence $s_a \le \varphi \le s_b$ a.e. in Ω. Furthermore, by the assertions of Lem. 2, the corresponding solutions $\{(\boldsymbol{u}_{n_j}, p_{n_j})\}_{j\in\mathbb{N}}$ satisfy $\boldsymbol{u}_{n_j} \to \boldsymbol{u}$ in $H^1(\Omega)$ and $p_{n_j} \to p$ in $L^2(\Omega)$ where $(\boldsymbol{u}, p) \in S_\varepsilon(\varphi)$.

For each $1 \le i \le m_1 + m_2$, by the continuity of $\boldsymbol{L}_i$ with respect to its second and third variables, it holds that $\boldsymbol{L}_i(x, \nabla\boldsymbol{u}_{n_j}, p_{n_j}) \to \boldsymbol{L}_i(x, \nabla\boldsymbol{u}, p)$ a.e. in Ω. Using the growth conditions in (A6), the strong convergences for $\{\boldsymbol{u}_{n_j}, p_{n_j}\}_{j\in\mathbb{N}}$ and the generalized Lebesgue dominated convergence theorem leads to

$$\boldsymbol{L}_i(x, \nabla\boldsymbol{u}_{n_j}, p_{n_j}) \to \boldsymbol{L}_i(x, \nabla\boldsymbol{u}, p) \text{ strongly in } L^2(\Omega) \text{ as } j \to \infty. \tag{27}$$

Together with the weak convergence $\nabla\varphi_{n_j}$ to $\nabla\varphi$ in $L^2(\Omega)$, we have

$$\lim_{j\to\infty} \int_\Omega \frac{1}{2}\nabla\varphi_{n_j} \cdot \boldsymbol{L}_i(x, \nabla\boldsymbol{u}_{n_j}, p_{n_j})\,\mathrm{dx} = \int_\Omega \frac{1}{2}\nabla\varphi \cdot \boldsymbol{L}_i(x, \nabla\boldsymbol{u}, p)\,\mathrm{dx}.$$

Note that $s_a \le \varphi_{n_j}, \varphi \le s_b$ a.e. in Ω for all $j \in \mathbb{N}$, and thus there exists a constant $M > 0$ such that $\sup_{x\in\Omega}\left(\left|\, y_i(x, \varphi_{n_j})\,\right|, |\, y_i(x, \varphi)\,|\right) \le M$ for all $n \in \mathbb{N}$. Using the splitting

$$\begin{aligned}
&\left| \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}_{n_j}, \nabla\boldsymbol{u}_{n_j}, p_{n_j})\ y_i(x, \varphi_{n_j}) - \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\ y_i(x, \varphi)\,\mathrm{dx} \right| \\
&\quad \le \left| \int_\Omega \left(\mathcal{K}_i(x, \boldsymbol{u}_{n_j}, \nabla\boldsymbol{u}_{n_j}, p_{n_j}) - \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\right)\ y_i(x, \varphi_{n_j})\,\mathrm{dx} \right| \\
&\quad + \left| \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\ (y_i(x, \varphi_{n_j}) - y_i(x, \varphi))\,\mathrm{dx} \right| =: I_1 + I_2,
\end{aligned}$$

we can show that $\lim_{n\to\infty} G_i(\varphi_{n_j}, \boldsymbol{u}_{n_j}, p_{n_j}) = G_i(\varphi, \boldsymbol{u}, p)$ once we demonstrate that $I_1, I_2 \to 0$ as $n \to \infty$. This would then imply that $\varphi \in \mathbb{K}_{ad}$. Using the growth conditions in (A5) for $\mathcal{K}_i$, the strong convergences for $\{(\boldsymbol{u}_{n_j}, p_{n_j})\}_{j\in\mathbb{N}}$ and the generalized Lebesgue dominated convergence theorem yields that

$$\mathcal{K}_i(x, \boldsymbol{u}_{n_j}, \nabla\boldsymbol{u}_{n_j}, p_{n_j}) \to \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p) \text{ strongly in } L^1(\Omega) \text{ as } j \to \infty.$$

Then, the assertion that $I_1 \to 0$ as $j \to \infty$ follows from the above strong convergence in $L^1(\Omega)$ and the boundedness of $y_i(x, \varphi_{n_j})$ in $L^\infty(\Omega)$. Meanwhile, dominating the sequence $\{\mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\ \ y_i(x, \varphi_{n_j})\}_{j\in\mathbb{N}}$ by the function $\|z_5\|_{L^\infty(\Omega)} M\,|\,\mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\,| \in L^1(\Omega)$, and the application of the usual Lebesgue dominating convergence theorem yields

$$\lim_{j\to\infty} \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\ y_i(x, \varphi_{n_j})\,\mathrm{dx} = \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p)\ y_i(x, \varphi)\,\mathrm{dx},$$

and hence $I_2 \to 0$ as $n \to \infty$. □

We state the existence result for a minimizer of the problem (17)-(19).

Theorem 1 *Under Assumptions* (A0)-(A8)*, there exists at least one minimizer to the problem* (17)-(19)*.*

Proof By (A8), $(\mathcal{B}+\mathcal{H})|_{\mathbb{K}_{ad}\times \boldsymbol{H}^1_{g,\sigma}(\Omega)\times L^2_0(\Omega)}$ is bounded from below by a constant $C_0 \in \mathbb{R}$. Then, by the non-negativity of $\hat{\alpha}_\varepsilon$ and Ψ, we find that there exists a constant $C_1 \in \mathbb{R}$ such that $\mathcal{J}_\varepsilon : \mathbb{K}_{ad} \times \boldsymbol{H}^1_{g,\sigma}(\Omega) \times L^2_0(\Omega) \to \mathbb{R}$ is bounded from below by C_1. Thus, we can choose a minimizing sequence $(\varphi_n, \boldsymbol{u}_n, p_n)_{n\in\mathbb{N}} \subset \mathbb{K}_{ad} \times \boldsymbol{H}^1_{g,\sigma}(\Omega) \times L^2_0(\Omega)$ such that $(\boldsymbol{u}_n, p_n) \in S_\varepsilon(\varphi_n)$ for all $n \in \mathbb{N}$ and

$$\lim_{n\to\infty} \mathcal{J}_\varepsilon(\varphi_n, \boldsymbol{u}_n, p_n) = \inf_{\varphi\in\mathbb{K}_{ad},(\boldsymbol{u},p)\in S_\varepsilon(\varphi)} \mathcal{J}_\varepsilon(\varphi, \boldsymbol{u}, p) \geq C_1 > -\infty.$$

Then, for arbitrary $\eta > 0$, there exists $N \in \mathbb{N}$ such that for $n > N$,

$$C_0 + \frac{\gamma\varepsilon}{4c_0}\|\nabla\varphi_n\|^2_{L^2(\Omega)} \leq \mathcal{J}_\varepsilon(\varphi_n, \boldsymbol{u}_n, p_n) \leq \inf_{\varphi\in\mathbb{K}_{ad},(\boldsymbol{u},p)\in S_\varepsilon(\varphi)} \mathcal{J}_\varepsilon(\varphi, \boldsymbol{u}, p) + \eta.$$

The above estimate implies that $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathbb{K}_{ad}$ is bounded uniformly in $H^1(\Omega) \cap L^\infty(\Omega)$. Thus, we may choose a subsequence $(\varphi_{n_k})_{k\in\mathbb{N}}$ such that $\varphi_{n_k} \to \varphi$ strongly in $L^p(\Omega)$ and almost everywhere in Ω for $2 \leq p < \infty$ in two-dimensions and $2 \leq p < 6$ in three-dimensions. Furthermore, by Lem. 5 we also have that $\varphi \in \mathbb{K}_{ad}$, and by Lem. 2, there is a subsequence $(\boldsymbol{u}_{n_k}, p_{n_k})_{n\in\mathbb{N}} \subset \boldsymbol{H}^1_{g,\sigma}(\Omega) \times L^2_0(\Omega)$ such that

$$\lim_{k\to\infty} \|\boldsymbol{u}_{n_k} - \boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega)} = 0, \quad \lim_{k\to\infty} \|p_{n_k} - p\|_{L^2(\Omega)} = 0,$$

for some $(\boldsymbol{u}, p) \in S_\varepsilon(\varphi)$, and

$$\lim_{k\to\infty} \int_\Omega \alpha_\varepsilon(\varphi_{n_k}) \left|\, \boldsymbol{u}_{n_k} \,\right|^2 \,\mathrm{dx} = \int_\Omega \alpha_\varepsilon(\varphi) \,|\, \boldsymbol{u} \,|^2 \,\mathrm{dx}\,.$$

The continuity of Ψ together with the fact that $(\varphi_{n_k})_{k\in\mathbb{N}} \subset L^\infty(\Omega)$ implies $(\Psi(\varphi_{n_k}))_{k\in\mathbb{N}}$ is a bounded sequence in $L^\infty(\Omega)$. The application of the dominated convergence theorem yields that $\Psi(\varphi_{n_k})$ converges strongly to $\Psi(\varphi)$ in $L^1(\Omega)$ as $k \to \infty$. Furthermore, by the weak lower semicontinuity assumptions of $\mathcal{B}$ and $\mathcal{H}$, and the weak lower semicontinuity of the mapping $\varphi \mapsto \|\nabla\varphi\|^2_{L^2(\Omega)}$, we find that

$$\mathcal{J}_\varepsilon(\varphi, \boldsymbol{u}, p) \leq \liminf_{k\to\infty} \mathcal{J}_\varepsilon(\varphi_{n_k}, \boldsymbol{u}_{n_k}, p_{n_k}) = \inf_{\phi\in\mathbb{K}_{ad},(v,q)\in S_\varepsilon(\phi)} \mathcal{J}_\varepsilon(\phi, v, q),$$

and so $(\varphi, \boldsymbol{u}, p) \in \mathbb{K}_{ad} \times \boldsymbol{H}^1_{g,\sigma}(\Omega) \times L^2_0(\Omega)$ is a minimizer of (17)-(19). □

From this point onwards, for fixed $\varepsilon > 0$, we denote a minimizer to the optimal control problem (17)-(19) as φ_ε with corresponding unique solution $(\boldsymbol{u}_\varepsilon, p_\varepsilon)$ to the state Eq. 18.

4 Optimality conditions

We use the notation $\mathrm{D}_j f$ to denote the partial derivative of f with respect to its jth variable. Furthermore, the notation $\left|\, \mathrm{D}_{(i,j)} f \,\right| \leq P$ means that the partial derivatives

$\mathrm{D}_i f$ and $\mathrm{D}_j f$ satisfy $|\,\mathrm{D}_i f\,| \le P$ and $\big|\,\mathrm{D}_j f\,\big| \le P$. To obtain optimality conditions, we make the following assumptions on the differentiability of B, z, h, $\mathcal{K}_i$, y_i, and $\boldsymbol{L}_i$.

(B1) In addition to (A3) assume further that $x \mapsto B(x, \boldsymbol{w}, \boldsymbol{A}, s)$, $x \mapsto z(x,t)$ and $x \mapsto h(x, \boldsymbol{A}, s, \boldsymbol{w})$ belong to $W^{1,1}(\Omega)$ for all $\boldsymbol{w} \in \mathbb{R}^d$, $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, $s, t \in \mathbb{R}$, and the partial derivatives

$$\begin{gathered}\mathrm{D}_2 B(x,\cdot,\boldsymbol{A},s),\ \mathrm{D}_3 B(x,\boldsymbol{w},\cdot,s),\ \mathrm{D}_4 B(x,\boldsymbol{w},\boldsymbol{A},\cdot),\ \mathrm{D}_2 z(x,\cdot),\\ \mathrm{D}_2 h(x,\cdot,s,\boldsymbol{w}),\ \mathrm{D}_3 h(x,\boldsymbol{A},\cdot,\boldsymbol{w}),\ \mathrm{D}_4 h(x,\boldsymbol{A},s,\cdot)\end{gathered}$$

exist for all $\boldsymbol{w} \in \mathbb{R}^d$, $s \in \mathbb{R}$, $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, and a.e. $x \in \Omega$ as Carathéodory functions with

$$\begin{aligned}|\,\mathrm{D}_2 B(x,\boldsymbol{w},\boldsymbol{A},s)\,| &\le \tilde{c}(x) + \tilde{b}_1(x)\,|\,\boldsymbol{w}\,|^{p-1} + \tilde{b}_2(x)\,|\,\boldsymbol{A}\,| + \tilde{b}_3(x)\,|\,s\,|,\\ \big|\,\mathrm{D}_{(3,4)} B(x,\boldsymbol{w},\boldsymbol{A},s)\,\big| &\le \tilde{a}(x) + \tilde{b}_1(x)\,|\,\boldsymbol{w}\,|^{p/2} + \tilde{b}_2(x)\,|\,\boldsymbol{A}\,| + \tilde{b}_3(x)\,|\,s\,|,\\ |\,\mathrm{D}_2 z(x,t)\,| &\le \tilde{b}_1(x),\\ \big|\,\mathrm{D}_{(2,3,4,5)} h(x,\boldsymbol{A},s,\boldsymbol{w})\,\big| &\le \tilde{a}(x) + \tilde{b}_1(x)\,|\,\boldsymbol{A}\,| + \tilde{b}_2(x)\,|\,s\,| + \tilde{b}_3(x)\,|\,\boldsymbol{w}\,|, \end{aligned} \tag{28}$$

for some non-negative functions $\tilde{a} \in L^2(\Omega)$, $\tilde{c} \in L^{\frac{p}{p-1}}(\Omega)$, $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3 \in L^\infty(\Omega)$, where $p \ge 2$ in two dimensions and $p \in [2,6]$ in three dimensions.

(B2) For each $1 \le i \le m_1 + m_2$, in addition to (A5) assume further that $x \mapsto \mathcal{K}_i(x, \boldsymbol{w}, \boldsymbol{A}, s)$, $x \mapsto k_i(x)$, $x \mapsto y_i(x,t)$, and $x \mapsto \boldsymbol{L}_i(x, \boldsymbol{A}, s)$ belong to $W^{1,1}(\Omega)$ for all $\boldsymbol{w} \in \mathbb{R}^d$, $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, $s, t \in \mathbb{R}$ and the partial derivatives

$$\begin{gathered}\mathrm{D}_2 \mathcal{K}_i(x,\cdot,\boldsymbol{A},s),\ \mathrm{D}_3 \mathcal{K}_i(x,\boldsymbol{w},\cdot,s),\ \mathrm{D}_4 \mathcal{K}_i(x,\boldsymbol{w},\boldsymbol{A},\cdot),\\ \mathrm{D}_2 y(x,\cdot),\ \mathrm{D}_2 \boldsymbol{L}_1(x,\cdot,s),\ \mathrm{D}_3 \boldsymbol{L}_1(x,\boldsymbol{A},\cdot)\end{gathered}$$

exist for all $\boldsymbol{w} \in \mathbb{R}^d$, $s \in \mathbb{R}$, $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, and a.e. $x \in \Omega$ as Carathéodory functions. Moreover, we assume that

$$\begin{aligned}|\,\mathrm{D}_2 \mathcal{K}_i(x,\boldsymbol{w},\boldsymbol{A},s)\,| &\le \tilde{c}(x) + \tilde{b}_1(x)\,|\,\boldsymbol{w}\,|^{p-1} + \tilde{b}_2(x)\,|\,\boldsymbol{A}\,| + \tilde{b}_3(x)\,|\,s\,|,\\ \big|\,\mathrm{D}_{(3,4)} \mathcal{K}_i(x,\boldsymbol{w},\boldsymbol{A},s)\,\big| &\le \tilde{a}(x) + \tilde{b}_1(x)\,|\,\boldsymbol{w}\,|^{p/2} + \tilde{b}_2(x)\,|\,\boldsymbol{A}\,| + \tilde{b}_3(x)\,|\,s\,|,\\ \big|\,\mathrm{D}_{(2,3)} \boldsymbol{L}_i(x,\boldsymbol{A},s)\,\big| &\le \tilde{b}_1(x),\\ |\,\mathrm{D}_2 y_i(x,t)\,| &\le \tilde{b}_1(x),\end{aligned}$$

for some non-negative functions $\tilde{a} \in L^2(\Omega)$, $\tilde{c} \in L^{\frac{p}{p-1}}(\Omega)$ and $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3 \in L^\infty(\Omega)$, where $p \ge 2$ in two dimensions and $p \in [2,6]$ in three dimensions.

Under (B1) and using [37, §4.3.3] or [17, Thm. 1 and 3], the Nemytskii operators

$$\begin{aligned}(L^2(\Omega))^{d\times d} \ni \boldsymbol{A} &\mapsto \mathrm{D}_2 h(\cdot,\boldsymbol{A},s,\boldsymbol{w}) \in L^2(\Omega) \quad \forall s \in L^2(\Omega), \boldsymbol{w} \in (L^2(\Omega))^d,\\ L^2(\Omega) \ni s &\mapsto \mathrm{D}_3 h(\cdot,\boldsymbol{A},s,\boldsymbol{w}) \in L^2(\Omega) \quad \forall \boldsymbol{A} \in (L^2(\Omega))^{d\times d}, \boldsymbol{w} \in (L^2(\Omega))^d,\\ (L^2(\Omega))^d \ni \boldsymbol{w} &\mapsto \mathrm{D}_4 h(\cdot,\boldsymbol{A},s,\boldsymbol{w}) \in L^2(\Omega) \quad \forall \boldsymbol{A} \in (L^2(\Omega))^{d\times d}, s \in L^2(\Omega),\end{aligned}$$

are well-defined and the operator

$$(L^2(\Omega))^{d\times d} \times L^2(\Omega) \times (L^2(\Omega))^d \ni (\boldsymbol{A}, s, \boldsymbol{w}) \mapsto h(\cdot,\boldsymbol{A},s,\boldsymbol{w}) \in L^1(\Omega)$$

is continuously Fréchet differentiable (see [17, Thm. 7] or [37, §4.3.3] with $p = r = 2$ and $q = 1$). Hence, we find that

$$\begin{aligned}\mathcal{H} : \left(H^1(\Omega) \cap L^\infty(\Omega)\right) \times \boldsymbol{H^1}(\boldsymbol{\Omega}) \times L^2(\Omega) &\to \mathbb{R} \\ (\varphi, \boldsymbol{u}, p) &\mapsto \int_\Omega \frac{1}{2} h(x, \nabla \boldsymbol{u}, p, \nabla \varphi)\,\mathrm{dx}\end{aligned}$$

is continuously Fréchet differentiable with derivative at $(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon)$ in the direction (η, v, s) given as

$$\begin{aligned}&\mathrm{D}\mathcal{H}(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon)(\eta, v, s) \\ &\quad = \textstyle\int_\Omega \frac{1}{2} (\mathrm{D}_2 h, \mathrm{D}_3 h, \mathrm{D}_4 h)\,|_{(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla \varphi_\varepsilon)} \cdot (\nabla v, s, \nabla \eta)\,\mathrm{dx}\,.\end{aligned} \tag{29}$$

Here we use the notation

$$\begin{aligned}&(\mathrm{D}_2 h, \mathrm{D}_3 h, \mathrm{D}_4 h)\,|_{(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla \varphi_\varepsilon)} \cdot (\nabla v, s, \nabla \eta) \\ &\quad := (\mathrm{D}_2 h) : \nabla v + (\mathrm{D}_3 h) s + (\mathrm{D}_4 h) \cdot \nabla \eta,\end{aligned}$$

where the partial derivatives are evaluated at $(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla \varphi_\varepsilon)$. With a similar argument, the mappings

$$\begin{aligned}\mathcal{B} : (H^1(\Omega) \cap L^\infty(\Omega)) \times \boldsymbol{H}^1(\Omega) \times L^2(\Omega) &\to \mathbb{R} \\ (\varphi, \boldsymbol{u}, p) &\mapsto \int_\Omega B(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p)\; z(x, \varphi)\,\mathrm{dx}\,, \\ G_i : (H^1(\Omega) \cap L^\infty(\Omega)) \times \boldsymbol{H}^1(\Omega) \times L^2(\Omega) &\to \mathbb{R} \\ (\varphi, \boldsymbol{u}, p) &\mapsto \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p)\; y_i(x, \varphi)\,\mathrm{dx} \\ &\quad + \int_\Omega k_i(x) + \frac{1}{2} \nabla \varphi \cdot \boldsymbol{L}_i(x, \nabla \boldsymbol{u}, p)\,\mathrm{dx}\,,\end{aligned}$$

for $1 \le i \le m_1 + m_2$, are continuously Fréchet differentiable, with derivatives at $(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon)$ in the direction (η, v, s) given as

$$\begin{aligned}&\mathrm{D}\mathcal{B}(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon)(\eta, v, s) \\ &\quad = \int_\Omega z(x, \varphi_\varepsilon)\; (\mathrm{D}_2 B, \mathrm{D}_3 B, \mathrm{D}_4 B)|_{(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)} \cdot (v, \nabla v, s)\,\mathrm{dx} \\ &\qquad + \int_\Omega B(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p)\; \mathrm{D}_2 z(x, \varphi_\varepsilon) \eta\,\mathrm{dx}\,,\end{aligned} \tag{30}$$

$$\begin{aligned}&\mathrm{D}G_i(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon)(\eta, v, s) \\ &\quad = \int_\Omega y_i(x, \varphi_\varepsilon)\; (\mathrm{D}_2 \mathcal{K}_i, \mathrm{D}_3 \mathcal{K}_i, \mathrm{D}_4 \mathcal{K}_i)|_{(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)} \cdot (v, \nabla v, s)\,\mathrm{dx} \\ &\qquad + \int_\Omega \mathcal{K}_i(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)\; \mathrm{D}_2 y_i(x, \varphi_\varepsilon) \eta\,\mathrm{dx} \\ &\qquad + \frac{1}{2} \int_\Omega \nabla \eta \cdot \boldsymbol{L}_i(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon) + \nabla \varphi_\varepsilon \cdot \left((\mathrm{D}_2 \boldsymbol{L}_i, \mathrm{D}_3 \boldsymbol{L}_i)|_{(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)} \cdot (\nabla v, s)\right)\,\mathrm{dx}\,.\end{aligned} \tag{31}$$

4.1 Fréchet differentiability of the objective functional

Due to the well-posedness of the state equations, we may now write the problem (17)-(19) as a minimizing problem for a reduced objective functional defined on an open set in $H^1(\Omega)\cap L^\infty(\Omega)$ with the help of Lem. 4. Let $(\varphi_\varepsilon, \boldsymbol{u}_\varepsilon, p_\varepsilon) \in \mathbb{K}_{ad} \times \boldsymbol{H}^1_{g,\sigma}(\Omega) \times L^2_0(\Omega)$ denote a minimizer of (17)-(19), obtained from Thm. 1. By Lem. 4, there exists a neighborhood $N \subset H^1(\Omega) \cap L^\infty(\Omega)$ of φ_ε such that for every $\psi \in N$, (18) is uniquely solvable. We define the reduced functional $j_\varepsilon : N \to \mathbb{R}$ by

$$j_\varepsilon(\psi) := \mathcal{J}_\varepsilon(\psi, S_\varepsilon(\psi)) \text{ for all } \psi \in N.$$

We now show that, as a mapping from $N \subset H^1(\Omega) \cap L^\infty(\Omega) \to \mathbb{R}$, j_ε is Fréchet differentiable at φ_ε. As Lem. 4 guarantees the Fréchet differentiability of the solution operator $S_\varepsilon(\varphi_\varepsilon)$ as a mapping from N to $\boldsymbol{H}^1(\Omega)\times L^2(\Omega)$, we focus on the dependence of $\mathcal{J}_\varepsilon$ on the first variable.

Fix $\varphi \in H^1(\Omega)$, then by (A2), $\hat{\alpha}_\varepsilon$ and $\hat{\alpha}'_\varepsilon$ are uniformly bounded and so

$$L^6(\Omega) \ni q \mapsto \hat{\alpha}'_\varepsilon(\varphi)q \in L^6(\Omega)$$

is a well-defined mapping from $H^1(\Omega) \subset L^6(\Omega)$ to $L^6(\Omega)$. By [37, §4.3.3], we see that $\hat{\alpha}_\varepsilon$ defines a Fréchet differentiable Nemytskii operator as a mapping from $L^6(\Omega)$ to $L^3(\Omega)$. Meanwhile, the assumption $\Psi \in C^{1,1}(\mathbb{R})$ and [37, Lem. 4.12] imply that $\Psi(\varphi)$ is continuously Fréchet differentiable Nemytskii operator as a mapping from $L^\infty(\Omega)$ to $L^\infty(\Omega)$. Combined with the Fréchet differentiability of the mapping $H^1(\Omega) \ni \varphi \mapsto \int_\Omega |\nabla\varphi|^2 \,\mathrm{dx}$, $\mathcal{B}$ and $\mathcal{H}$, we obtain that $j_\varepsilon : N \to \mathbb{R}$ is Fréchet differentiable.

4.2 Existence of Lagrange multipliers

To show the existence of Lagrange multipliers for the integral constraints, we make use of the Zowe–Kurcyusz constraint qualification (ZKCQ), see [39] and [37, §6.1.2] for more details. For this purpose, we introduce the notation

$$\begin{aligned}
\mathbb{Y} &:= \mathbb{R}^{m_1+m_2},\\
\mathbb{K} &:= \left\{y \in Y \mid y_i = 0,\, y_j \geq 0 \text{ for} 1 \leq i \leq m_1,\, m_1+1 \leq j \leq m_1+m_2\right\} \subset \mathbb{Y},\\
\mathcal{G}_i(\varphi) &:= G_i(\varphi, S_\varepsilon(\varphi)) \text{ for } 1 \leq i \leq m_1+m_2,\\
g(\varphi) &:= \left(\mathcal{G}_1(\varphi), \ldots, \mathcal{G}_{m_1}(\varphi), \mathcal{G}_{m_1+1}(\varphi), \ldots, \mathcal{G}_{m_1+m_2}(\varphi)\right),
\end{aligned}$$

and recall the set

$$\Phi = \{f \in H^1(\Omega) \mid s_a \leq f \leq s_b \text{ a.e. in } \Omega\}.$$

Then, Φ is a closed convex subset of $H^1(\Omega)$ and $\mathbb{K}$ is a closed convex cone in $\mathbb{Y}$ with vertex at the origin, i.e., $\delta_1\mathbb{K} + \delta_2\mathbb{K} \subset \mathbb{K}$ for $\delta_1, \delta_2 > 0$. In the notation of [39], we introduce the sets

$$\begin{aligned}
\Phi(\varphi_\varepsilon) &= \{\beta(\varphi - \varphi_\varepsilon) \mid \varphi \in \Phi, \beta \geq 0\},\\
\mathbb{K}(g(\varphi_\varepsilon)) &= \{\eta - \beta g(\varphi_\varepsilon) \mid \eta \in \mathbb{K}, \beta \geq 0\}.
\end{aligned}$$

Fix $1 \leq i \leq m_1 + m_2$ and an arbitrary function $\zeta \in \Phi$. Convexity of Φ implies that $\varphi_\varepsilon + t(\zeta - \varphi_\varepsilon) \in \Phi$ for sufficiently small values of t. Then, denoting the linearized state variables associated to $\delta = \zeta - \varphi_\varepsilon$ as $(\boldsymbol{w}_\varepsilon, r_\varepsilon)$ (see Lem. 4), the mapping $\phi \mapsto \mathcal{G}_i(\phi) = G_i(\phi, S_\varepsilon(\phi))$ is continuously Fréchet differentiable at φ_ε with derivative in the direction $\zeta - \varphi_\varepsilon$ given as (see also (31))

$$\begin{aligned} &\mathrm{D}\mathcal{G}_i(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) \\ &\quad = \int_\Omega (\mathrm{D}_2 K_i, \mathrm{D}_3 K_i, \mathrm{D}_4 K_i, \mathrm{D}_5 K_i) \cdot (\boldsymbol{w}_\varepsilon, \nabla \boldsymbol{w}_\varepsilon, r_\varepsilon, \zeta - \varphi_\varepsilon)\,\mathrm{dx} \\ &\qquad + \tfrac{1}{2} \int_\Omega \nabla(\zeta - \varphi_\varepsilon) \cdot \boldsymbol{L}_i(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)\,\mathrm{dx} \\ &\qquad + \tfrac{1}{2} \int_\Omega \nabla \varphi_\varepsilon \cdot \left[(\mathrm{D}_2 \boldsymbol{L}_i, \mathrm{D}_3 \boldsymbol{L}_i)|_{(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon)} \cdot (\nabla \boldsymbol{w}_\varepsilon, r_\varepsilon) \right] \mathrm{dx}\,, \end{aligned} \tag{32}$$

where $\mathrm{D}_2 K_i$, $\mathrm{D}_3 K_i$, $\mathrm{D}_4 K_i$ and $\mathrm{D}_5 K_i$ are evaluated at $(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \varphi_\varepsilon)$. Then, it holds that

$$g'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) = (\mathrm{D}\mathcal{G}_1(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon), \ldots, \mathrm{D}\mathcal{G}_{m_1+m_2}(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon)).$$

The existence of bounded Lagrange multipliers $\lambda := (\lambda_1, \ldots, \lambda_{m_1+m_2}) \in \mathbb{K}^+ := \{y \in \mathbb{Y} \mid y \cdot \eta = 0\ \forall \eta \in \mathbb{K}\}$ satisfying

$$\lambda \cdot g(\varphi_\varepsilon) = 0, \text{ and } \langle \mathrm{D} j_\varepsilon(\varphi_\varepsilon) + \lambda \cdot g'(\varphi_\varepsilon), \zeta - \varphi_\varepsilon \rangle \geq 0 \quad \forall \zeta \in \Phi,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^1(\Omega)$ and its dual, follows if φ_ε is a regular point in the sense of [39], or equivalently the so-called Zowe–Kurcyusz constraint qualification

$$\mathbb{Y} = g'(\varphi_\varepsilon)\Phi(\varphi_\varepsilon) - \mathbb{K}(g(\varphi_\varepsilon)) \tag{33}$$

has to hold. We now make the following assumption:

(C1) For any $z \in \mathbb{Y} = \mathbb{R}^{m_1+m_2}$, there exists a function $\psi_* \in \Phi$, vectors $\tau \in \mathbb{Y}$, $\xi, \eta \in \mathbb{R}^{m_2}$ such that $\tau_i \geq 0$, $\xi_j \geq 0$, $\eta_j \geq 0$ for $1 \leq i \leq m_1 + m_2$ and $1 \leq j \leq m_2$, and

$$\begin{aligned} z_i &= \tau_i \mathrm{D}\mathcal{G}_i(\varphi_\varepsilon)(\psi_* - \varphi_\varepsilon), && \text{for } 1 \leq i \leq m_1 \\ z_{m_1+j} &= \tau_{m_1+j} \mathrm{D}\mathcal{G}_{m_1+j}(\varphi_\varepsilon)(\psi_* - \varphi_\varepsilon) - \eta_j + \xi_j \mathcal{G}_{m_1+j}(\varphi_\varepsilon), && \text{for } 1 \leq j \leq m_2. \end{aligned}$$

Then, under (C1) and using [39, Thm. 3.1 and 4.1] there exist $\lambda_1, \ldots, \lambda_{m_1} \in \mathbb{R}$, $\lambda_{m_1+1}, \ldots, \lambda_{m_1+m_2} \in \mathbb{R}_{\geq 0}$ such that

$$\begin{aligned} \mathrm{D} j_\varepsilon(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) &+ \sum_{i=1}^{m_1} \lambda_i \mathrm{D}\mathcal{G}_i(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) \\ &+ \sum_{j=1}^{m_2} \lambda_{m_1+j} \mathrm{D}\mathcal{G}_{m_1+j}(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) \geq 0 \quad \forall \zeta \in \Phi \end{aligned} \tag{34}$$

holds with the following complementary slackness conditions for the inequality constraints

$$\lambda_{m_1+j} \mathcal{G}_{m_1+j}(\varphi_\varepsilon) = 0 \text{ for } 1 \leq j \leq m_2. \tag{35}$$

We mention that (33) is equivalent (see [39, §3] and [21, Thm. 1.56]) to the following interior point/linearized Slater condition (which is also commonly known as the Robinson regularity condition [30]):

$$\mathbf{0} \in \operatorname{int}\left(g(\varphi_\varepsilon) + g'(\varphi_\varepsilon)(\Phi - \varphi_\varepsilon) - \mathbb{K}\right).$$

4.3 Adjoint system

We now introduce the Lagrangian $\mathbb{L} : (H^1(\Omega) \cap L^\infty(\Omega)) \times \boldsymbol{H}^1(\Omega) \times L^2(\Omega) \times \boldsymbol{H}^1_0(\Omega) \times L^2(\Omega) \to \mathbb{R}$ as

$$\begin{aligned}
&\mathbb{L}(\varphi, \boldsymbol{u}, p, \boldsymbol{q}, \pi) \\
&\quad := \int_\Omega \frac{1}{2}\hat{\alpha}_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2 + \frac{\gamma}{2c_0}\left(\frac{1}{\varepsilon}\Psi(\varphi) + \frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right) \mathrm{dx} \\
&\qquad + \int_\Omega b(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p, \varphi) + \frac{1}{2}h(x, \nabla\boldsymbol{u}, p, \nabla\varphi)\,\mathrm{dx} \\
&\qquad - \int_\Omega \alpha_\varepsilon(\varphi)\boldsymbol{u}\cdot\boldsymbol{q} + \mu\nabla\boldsymbol{u}\cdot\nabla\boldsymbol{q} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\boldsymbol{q} - p\,\mathrm{div}\,\boldsymbol{q} - \boldsymbol{f}\cdot\boldsymbol{q} - \pi\,\mathrm{div}\,\boldsymbol{u}\,\mathrm{dx} \\
&\qquad + \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i\left(K_i(x, \boldsymbol{u}, \nabla\boldsymbol{u}, p, \varphi) + \frac{1}{2}\nabla\varphi\cdot\boldsymbol{L}_i(x, \nabla\boldsymbol{u}, p)\right) + \theta p\,\mathrm{dx}
\end{aligned}$$

where λ_i is the Lagrange multiplier for the integral constraint $\mathcal{G}_i(\varphi)$ and θ is a Lagrange multiplier for the constraint $\int_\Omega p\,\mathrm{dx} = 0$ for the pressure. A formal computation of $\mathrm{D}_{\boldsymbol{u}}\mathbb{L}$ and $\mathrm{D}_p\mathbb{L}$ yields the following adjoint system for the minimizer φ_ε:

$$\begin{aligned}
\alpha_\varepsilon(\varphi_\varepsilon)\,\boldsymbol{q}_\varepsilon\; & -\mu\,\mathrm{div}\,(\nabla\boldsymbol{q}_\varepsilon + (\nabla\boldsymbol{q}_\varepsilon)^\top) + (\nabla\boldsymbol{u}_\varepsilon)^\top\boldsymbol{q}_\varepsilon - (\boldsymbol{u}_\varepsilon\cdot\nabla)\boldsymbol{q}_\varepsilon + \nabla\pi_\varepsilon \\
&= \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon + \mathrm{D}_2 b - \mathrm{div}\,(\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h) \\
&\quad + \sum_{i=1}^{m_1+m_2}\left(\lambda_i \mathrm{D}_2 K_i - \mathrm{div}\,\left(\lambda_i\left(\mathrm{D}_3 K_i + \tfrac{1}{2}\nabla\varphi_\varepsilon\cdot\mathrm{D}_2\boldsymbol{L}_i\right)\right)\right) \qquad \text{in } \Omega, \quad (36a) \\
\mathrm{div}\,\boldsymbol{q}_\varepsilon\; &= -\mathrm{D}_4 b - \tfrac{1}{2}\mathrm{D}_3 h - \theta - \sum_{i=1}^{m_1+m_2}\left(\lambda_i \mathrm{D}_4 L_i + \tfrac{1}{2}\lambda_i\nabla\varphi_\varepsilon\cdot\mathrm{D}_3\boldsymbol{L}_{1,i}\right) \quad \text{in } \Omega, \quad (36b) \\
\boldsymbol{q}_\varepsilon\; &= \mathbf{0} \qquad \text{on } \partial\Omega, \quad (36c)
\end{aligned}$$

where $\mathrm{D}_{(2,3,4)}b$ are evaluated at $(x, \boldsymbol{u}_\varepsilon, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon, \varphi_\varepsilon)$, $\mathrm{D}_{(2,3)}h$ are evaluated at $(x, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla\varphi_\varepsilon)$, $\mathrm{D}_{(2,3,4)}K_i$ are evaluated at $(x, \boldsymbol{u}_\varepsilon, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon, \varphi_\varepsilon)$, and $\mathrm{D}_{(2,3)}\boldsymbol{L}_i$ are evaluated at $(x, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon)$, and upon integrating the divergence equation for $\boldsymbol{q}_\varepsilon$, we obtain

$$\theta = \tfrac{1}{|\Omega|}\int_\Omega -\mathrm{D}_4 b - \tfrac{1}{2}\mathrm{D}_3 h - \sum_{i=1}^{m_1+m_2}\left(\lambda_i \mathrm{D}_4 K_i + \tfrac{1}{2}\lambda_i\nabla\varphi_\varepsilon\cdot\mathrm{D}_3\boldsymbol{L}_i\right)\mathrm{dx}. \qquad (37)$$

Let us also recall from (25) and (26) that

$$\begin{aligned} \mathrm{D}_{(2,3,4)} b(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi) &= z(x, \varphi)\, \mathrm{D}_{(2,3,4)} B(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p), \\ \mathrm{D}_{(2,3,4)} K_i(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi) &= y_i(x, \varphi)\, \mathrm{D}_{(2,3,4)} \mathcal{K}_i(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p). \end{aligned}$$

We now show that the adjoint system is well-posed.

Lemma 6 *Let $\varphi_\varepsilon \in H^1(\Omega) \cap L^\infty(\Omega)$ be the minimizer obtained from Thm. 1 and $(\boldsymbol{u}_\varepsilon, p_\varepsilon) = S_\varepsilon(\varphi_\varepsilon)$. Furthermore, let $\{\lambda_i\}_{i=1}^{m_1+m_2}$ be the Lagrange multipliers associated to the integral state constraints $\{\mathcal{G}_i(\varphi_\varepsilon)\}_{i=1}^{m_1+m_2}$. Then, under (B1), (B2) and (C1), there exists a unique weak solution pair $(\boldsymbol{q}_\varepsilon, \pi_\varepsilon) \in \boldsymbol{H}_0^1(\Omega) \times L^2(\Omega)$ to the adjoint system (36) in the following sense*

$$\begin{aligned} &\int_\Omega \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{q}_\varepsilon \cdot v + \mu(\nabla \boldsymbol{q}_\varepsilon + (\nabla \boldsymbol{q}_\varepsilon)^\top) \cdot \nabla v + (\nabla \boldsymbol{u}_\varepsilon)^\top \boldsymbol{q}_\varepsilon \cdot v - (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{q}_\varepsilon \cdot v \,\mathrm{dx} \\ &\quad = \int_\Omega \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot v + (\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h) \cdot \nabla v + \mathrm{D}_2 b \cdot v \,\mathrm{dx} \\ &\qquad + \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i (\mathrm{D}_2 K_i \cdot v + (\tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_2 \boldsymbol{L}_i + \mathrm{D}_3 K_i) \cdot \nabla v) \,\mathrm{dx} \end{aligned} \tag{38}$$

for all $v \in \boldsymbol{H}_{0,\sigma}^1(\Omega)$.

Proof For convenience, we use the notation

$$g := -\mathrm{D}_4 b - \tfrac{1}{2}\mathrm{D}_3 h - \theta - \sum_{i=1}^{m_1+m_2} \lambda_i \left(\mathrm{D}_4 K_i + \tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_3 \boldsymbol{L}_i \right), \tag{39}$$

so that (36b) reads as $\operatorname{div} \boldsymbol{q}_\varepsilon = g$ in Ω. Then, by (B1), (B2) and (37), we see that g belongs to the function space $L_0^2(\Omega)$.

Applying [33, Lem. II.2.1.1], we find a vector field $G \in \boldsymbol{H}_0^1(\Omega)$ such that

$$\operatorname{div} G = g \text{ in } \Omega, \text{ and } \|\nabla G\|_{L^2(\Omega)} \le C \|g\|_{L^2(\Omega)}$$

for some constant $C > 0$ depending only on Ω. We define the bilinear form $a : \boldsymbol{H}_{0,\sigma}^1(\Omega) \times \boldsymbol{H}_{0,\sigma}^1(\Omega) \to (\boldsymbol{H}_{0,\sigma}^1(\Omega))'$ by

$$\begin{aligned} a(z, v) := &\int_\Omega \alpha_\varepsilon(\varphi_\varepsilon) z \cdot v + \mu(\nabla z + (\nabla z)^\top) \cdot \nabla v \,\mathrm{dx} \\ &+ \int_\Omega (\nabla \boldsymbol{u}_\varepsilon)^\top z \cdot v - (\boldsymbol{u}_\varepsilon \cdot \nabla) z \cdot v \,\mathrm{dx}\,. \end{aligned} \tag{40}$$

Using (22), Poincaré's inequality, Hölder's inequality, the boundedness of α_ε and properties of the trilinear form $b(\boldsymbol{u}, v, \boldsymbol{w}) := \int_\Omega (\boldsymbol{u} \cdot \nabla) v \cdot \boldsymbol{w} \,\mathrm{dx}$ (see [14, Lem. 4.1]), it can be shown similar to [14, Proof of Lem. 4.9] that $a(\cdot, \cdot)$ is a bounded and coercive bilinear form. Furthermore, defining

$$\begin{aligned} F(v) := &\int_\Omega \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot v + (\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h) \cdot \nabla v + \mathrm{D}_2 b \cdot v \,\mathrm{dx} \\ &+ \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i \left(\mathrm{D}_2 K_i \cdot v + \left(\tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_2 \boldsymbol{L}_i + \mathrm{D}_3 K_i \right) \cdot \nabla v \right) \mathrm{dx} \\ &- \int_\Omega \alpha_\varepsilon(\varphi_\varepsilon) G \cdot v + \mu(\nabla G + (\nabla G)^\top) \cdot \nabla v \,\mathrm{dx} \\ &- \int_\Omega (\nabla \boldsymbol{u}_\varepsilon)^\top G \cdot v - (\boldsymbol{u}_\varepsilon \cdot \nabla) G \cdot v \,\mathrm{dx}\,, \end{aligned} \tag{41}$$

and applying (B1), (B2), the fact that $G, \boldsymbol{u}_\varepsilon \in \boldsymbol{H}^1(\Omega)$ and Sobolev embeddings leads to the deduction that $F(v)$ is a bounded linear form on $\boldsymbol{H}_{0,\sigma}^1(\Omega)$.

Thus, by the Lax–Milgram theorem, we obtain a unique $\hat{\boldsymbol{q}} \in \boldsymbol{H}^1_{0,\sigma}(\Omega)$ such that

$$a(\hat{\boldsymbol{q}}, v) = F(v).$$

This implies that the solution $\boldsymbol{q}_\varepsilon := \hat{\boldsymbol{q}} + G \in \boldsymbol{H}^1_0(\Omega)$ satisfies the weak formulation (38) with

$$\operatorname{div} \boldsymbol{q}_\varepsilon = \operatorname{div} G = g.$$

The existence of a unique adjoint pressure $\pi_\varepsilon \in L^2(\Omega)$ follows from standard results, see for instance [33, Lem. II.2.2.1]. Thus $(\boldsymbol{q}_\varepsilon, \pi_\varepsilon)$ is the unique weak solution to the adjoint system (36). □

4.4 Necessary optimality conditions

Now we can formulate the first order necessary optimality conditions for our optimal control problem.

Theorem 2 *Let $\varphi_\varepsilon \in \mathbb{K}_{ad}$ be a minimizer of (17)-(19) with corresponding (unique) state variables $(\boldsymbol{u}_\varepsilon, p_\varepsilon) = S_\varepsilon(\varphi_\varepsilon)$, $\boldsymbol{u}_\varepsilon \in \boldsymbol{H}^1_{g,\sigma}(\Omega)$, $p_\varepsilon \in L^2_0(\Omega)$. Furthermore, let $\{\lambda_i\}_{i=1}^{m_1+m_2}$ be the Lagrange multipliers associated to the integral state constraints $\{\mathcal{G}_i(\varphi_\varepsilon)\}_{i=1}^{m_1+m_2}$, and $(q_\varepsilon, \pi_\varepsilon)$ be the unique solution to the adjoint system (36). Then, under (A0)-(C1), the following optimality system is fulfilled:*

$$\begin{aligned} 0 \le{}& \left\langle \tfrac{1}{2}\hat{\alpha}'_\varepsilon(\varphi_\varepsilon) \,|\, \boldsymbol{u}_\varepsilon \,|^2 - \alpha'_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + \tfrac{\gamma}{2c_0\varepsilon}\Psi'(\varphi_\varepsilon), \zeta - \varphi_\varepsilon \right\rangle \\ &+ \left\langle \mathrm{D}_5 b + \sum_{i=1}^{m_1+m_2} \lambda_i \mathrm{D}_5 K_i, \zeta - \varphi_\varepsilon \right\rangle_{L^2(\Omega)} \\ &+ \left\langle \tfrac{\gamma\varepsilon}{2c_0}\nabla\varphi_\varepsilon + \tfrac{1}{2}\mathrm{D}_4 h + \tfrac{1}{2}\sum_{i=1}^{m_1+m_2} \lambda_i \boldsymbol{L}_i, \nabla(\zeta - \varphi_\varepsilon) \right\rangle_{L^2(\Omega)} \qquad \forall \zeta \in \Phi, \end{aligned} \tag{42}$$

where $\mathrm{D}_4 h$ is evaluated at $(x, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla\varphi_\varepsilon)$, $\boldsymbol{L}_i$ is evaluated at $(x, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon)$, and $\mathrm{D}_5 b = B(x, \boldsymbol{u}_\varepsilon, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon)\, \mathrm{D}_2 z(x, \varphi_\varepsilon)$, $\mathrm{D}_5 K_i = \mathcal{K}_i(x, \boldsymbol{u}_\varepsilon, \nabla\boldsymbol{u}_\varepsilon, p_\varepsilon)\, \mathrm{D}_2 y_i(x, \varphi_\varepsilon)$.

Proof In Section 4.1 we have shown that the reduced functional

$$j_\varepsilon(\varphi_\varepsilon) := \mathcal{J}_\varepsilon(\varphi_\varepsilon, S_\varepsilon(\varphi_\varepsilon))$$

is Fréchet differentiable with respect to φ_ε, and in Section 4.2 we derived the gradient Eq. 34. We now want to rewrite (34) into a more convenient form using the adjoint system. For $\zeta \in \Phi$, let $(\boldsymbol{w}_\varepsilon, r_\varepsilon)$ denote the unique solution to the linearized state Eqs. 24 corresponding to $\delta = \zeta - \varphi_\varepsilon$. Then, computing the derivative of j_ε at φ_ε in the direction δ leads to

$$\begin{aligned} &\mathrm{D} j_\varepsilon(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) \\ &\quad = \int_\Omega \tfrac{1}{2}\hat{\alpha}'_\varepsilon(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) \,|\, \boldsymbol{u}_\varepsilon \,|^2 + \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{w}_\varepsilon \,\mathrm{dx} \\ &\qquad + \int_\Omega \tfrac{\gamma}{2c_0}\left(\tfrac{1}{\varepsilon}\Psi'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon) + \varepsilon\nabla\varphi_\varepsilon \cdot \nabla(\zeta - \varphi_\varepsilon)\right) \mathrm{dx} \\ &\qquad + \int_\Omega (\mathrm{D}_2 b, \mathrm{D}_3 b, \mathrm{D}_4 b, \mathrm{D}_5 b) \cdot (\boldsymbol{w}_\varepsilon, \nabla\boldsymbol{w}_\varepsilon, r_\varepsilon, \zeta - \varphi_\varepsilon) \,\mathrm{dx} \\ &\qquad + \int_\Omega \tfrac{1}{2}(\mathrm{D}_2 h, \mathrm{D}_3 h, \mathrm{D}_4 h) \cdot (\nabla\boldsymbol{w}_\varepsilon, r_\varepsilon, \nabla(\zeta - \varphi_\varepsilon)) \,\mathrm{dx}\,, \end{aligned} \tag{43}$$

where in the above and for the rest of the proof $\{\mathrm{D}_i b\}_{i=2}^5$ are evaluated at $(x, \boldsymbol{u}_\varepsilon, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \varphi_\varepsilon)$ and $\{\mathrm{D}_i h\}_{i=2}^4$ are evaluated at $(x, \nabla \boldsymbol{u}_\varepsilon, p_\varepsilon, \nabla \varphi_\varepsilon)$. Using the adjoint state $\boldsymbol{q}_\varepsilon$ as a test function in (24) (with $\delta = \zeta - \varphi_\varepsilon$) leads to

$$\begin{aligned} 0 = &\int_\Omega \alpha_\varepsilon'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{w}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + \mu \nabla \boldsymbol{w}_\varepsilon \cdot \nabla \boldsymbol{q}_\varepsilon \,\mathrm{dx} \\ &+ \int_\Omega (\boldsymbol{w}_\varepsilon \cdot \nabla)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{w}_\varepsilon \cdot \boldsymbol{q}_\varepsilon - r_\varepsilon g \,\mathrm{dx}\,, \end{aligned} \tag{44}$$

where $g = \operatorname{div} q_\varepsilon$ as in (39). Using the linearized state $\boldsymbol{w}_\varepsilon$ as a test function in the adjoint system (38) leads to

$$\begin{aligned} &\int_\Omega \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + \mu \nabla \boldsymbol{q}_\varepsilon \cdot \nabla \boldsymbol{w}_\varepsilon + (\nabla \boldsymbol{u}_\varepsilon)^\top \boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon - (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon \,\mathrm{dx} \\ &\quad = \int_\Omega \hat{\alpha}(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + (\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h) \cdot \nabla \boldsymbol{w}_\varepsilon + \mathrm{D}_2 b \cdot \boldsymbol{w}_\varepsilon \,\mathrm{dx} \\ &\qquad + \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i (\mathrm{D}_2 K_i \cdot \boldsymbol{w}_\varepsilon + (\tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_2 \boldsymbol{L}_i + \mathrm{D}_3 K_i) \cdot \nabla \boldsymbol{w}_\varepsilon) \,\mathrm{dx}\,, \end{aligned} \tag{45}$$

where we have used $\boldsymbol{w}_\varepsilon \in \boldsymbol{H}^1_{0,\sigma}(\Omega)$ to deduce that $\int_\Omega (\nabla q_\varepsilon)^\top \cdot \nabla \boldsymbol{w}_\varepsilon \,\mathrm{dx} = 0$ (see for instance [14, (4.29)]). Upon comparing terms in (44) and (45) we find that

$$\begin{aligned} &\int_\Omega \hat{\alpha}(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + (\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h) \cdot \nabla \boldsymbol{w}_\varepsilon + \mathrm{D}_2 b \cdot \boldsymbol{w}_\varepsilon + (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon \,\mathrm{dx} \\ &\quad + \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i (\mathrm{D}_2 K_i \cdot \boldsymbol{w}_\varepsilon + (\tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_2 \boldsymbol{L}_i + \mathrm{D}_3 K_i) \cdot \nabla \boldsymbol{w}_\varepsilon) \,\mathrm{dx} \\ &= \int_\Omega \alpha_\varepsilon(\varphi_\varepsilon)\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + \mu \nabla q_\varepsilon \cdot \nabla \boldsymbol{w}_\varepsilon + (\boldsymbol{w}_\varepsilon \cdot \nabla)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon \,\mathrm{dx} \\ &= \int_\Omega r_\varepsilon g - \alpha_\varepsilon'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon - (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{w}_\varepsilon \cdot \boldsymbol{q}_\varepsilon \end{aligned} \tag{46}$$

Using that $r_\varepsilon \in L^2_0(\Omega)$, $\operatorname{div} \boldsymbol{u}_\varepsilon = 0$ in Ω, $\boldsymbol{q}_\varepsilon = \boldsymbol{w}_\varepsilon = \boldsymbol{0}$ on $\partial\Omega$, and thus

$$\int_\Omega r_\varepsilon \theta \,\mathrm{dx} = \theta \int_\Omega r_\varepsilon \,\mathrm{dx} = 0,$$

$$\int_\Omega (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + (\boldsymbol{u}_\varepsilon \cdot \nabla)\boldsymbol{w}_\varepsilon \cdot \boldsymbol{q}_\varepsilon \,\mathrm{dx} = \int_\Omega \boldsymbol{u}_\varepsilon \cdot \nabla(\boldsymbol{q}_\varepsilon \cdot \boldsymbol{w}_\varepsilon) \,\mathrm{dx} = 0,$$

we can simplify (46) into

$$\begin{aligned} &\int_\Omega r_\varepsilon \left(-\mathrm{D}_4 b - \tfrac{1}{2}\mathrm{D}_3 h - \sum_{i=1}^{m_1+m_2} \lambda_i \left(\mathrm{D}_4 K_i + \tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_3 \boldsymbol{L}_i \right) \right) \mathrm{dx} \\ &\qquad - \int_\Omega \alpha_\varepsilon'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon \,\mathrm{dx} \\ &= \int_\Omega \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + \left(\mathrm{D}_3 b + \tfrac{1}{2}\mathrm{D}_2 h \right) \cdot \nabla \boldsymbol{w}_\varepsilon + \mathrm{D}_2 b \cdot \boldsymbol{w}_\varepsilon \,\mathrm{dx} \\ &\quad + \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i \left(\mathrm{D}_2 K_i \cdot \boldsymbol{u} + \left(\tfrac{1}{2}\nabla \varphi_\varepsilon \cdot \mathrm{D}_2 \boldsymbol{L}_i + \mathrm{D}_3 K_i \right) \cdot \nabla \boldsymbol{w}_\varepsilon \right) \mathrm{dx}\,, \end{aligned}$$

and upon rearranging we obtain

$$\begin{aligned} &\int_\Omega \hat{\alpha}_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{w}_\varepsilon + (\mathrm{D}_2 b, \mathrm{D}_3 b, \mathrm{D}_4 b) \cdot (\boldsymbol{w}_\varepsilon, \nabla \boldsymbol{w}_\varepsilon, r_\varepsilon) \,\mathrm{dx} \\ &\quad + \int_\Omega \tfrac{1}{2}(\mathrm{D}_2 h, \mathrm{D}_3 h) \cdot (\nabla \boldsymbol{w}_\varepsilon, r_\varepsilon) \,\mathrm{dx} \\ &\quad = \int_\Omega -\alpha_\varepsilon'(\varphi_\varepsilon)(\zeta - \varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon \,\mathrm{dx} \\ &\qquad - \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i (\mathrm{D}_2 K_i, \mathrm{D}_3 K_i, \mathrm{D}_4 K_i) \cdot (\boldsymbol{w}_\varepsilon, \nabla \boldsymbol{w}_\varepsilon, r_\varepsilon) \,\mathrm{dx} \\ &\qquad - \int_\Omega \sum_{i=1}^{m_1+m_2} \lambda_i \tfrac{1}{2}\nabla \varphi_\varepsilon \cdot (\mathrm{D}_2 \boldsymbol{L}_i, \mathrm{D}_3 \boldsymbol{L}_i) \cdot (\nabla \boldsymbol{w}_\varepsilon, r_\varepsilon) \,\mathrm{dx}\,. \end{aligned} \tag{47}$$

Substituting (47) into (43), we obtain

$$\begin{aligned}
&\mathrm{D}j_\varepsilon(\varphi_\varepsilon)(\zeta-\varphi_\varepsilon)\\
&\quad=\int_\Omega\left(\tfrac{1}{2}\hat{\alpha}'_\varepsilon(\varphi_\varepsilon)\,|\,\boldsymbol{u}_\varepsilon\,|^2-\alpha'_\varepsilon(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon\cdot\boldsymbol{q}_\varepsilon+\tfrac{\gamma}{2c_0\varepsilon}\Psi'(\varphi_\varepsilon)\right)(\zeta-\varphi_\varepsilon)\,\mathrm{d}x\\
&\qquad+\int_\Omega \mathrm{D}_5 b\,(\zeta-\varphi_\varepsilon)+\left(\tfrac{\gamma}{2c_0}\varepsilon\nabla\varphi_\varepsilon+\tfrac{1}{2}\mathrm{D}_4 h\right)\cdot\nabla(\zeta-\varphi_\varepsilon)\,\mathrm{d}x\\
&\qquad-\int_\Omega\sum_{i=1}^{m_1+m_2}\lambda_i\,(\mathrm{D}_2K_i,\mathrm{D}_3K_i,\mathrm{D}_4K_i)\cdot(\boldsymbol{w}_\varepsilon,\nabla\boldsymbol{w}_\varepsilon,r_\varepsilon)\,\mathrm{d}x\\
&\qquad-\int_\Omega\sum_{i=1}^{m_1+m_2}\lambda_i\,\tfrac{1}{2}\nabla\varphi_\varepsilon\cdot(\mathrm{D}_2\boldsymbol{L}_i,\mathrm{D}_3\boldsymbol{L}_i)\cdot(\nabla\boldsymbol{w}_\varepsilon,r_\varepsilon)\,\mathrm{d}x\,.
\end{aligned}\tag{48}$$

Together with the gradient Eq. 34 and the distributional derivatives (32), we then obtain (42). □

Remark 1 In the case where there is only a volume constraint, i.e., $m_1+m_2=1$ with $\mathcal{G}(\varphi):=\int_\Omega\varphi-\beta\,\mathrm{d}x$ for a fixed constant $\beta\in(-1,1)$, the existence of Lagrange multipliers using the Zowe–Kurcyusz constraint qualification has been shown in [18, Proof of Thm. 7.1] (for the case of inequality constraint), see also [12, Proof of Thm. 3] for another argument using geometric variations. For the case of equality constraint, we refer to [14, Proof of Thm. 4.10] which is based on a different argument.

5 Verification of constraint qualification

In this section, we consider a model problem of minimizing the drag subject to constraints on the mass, center of mass and volume of the object. More precisely, in a bounded domain $\Omega\subset\mathbb{R}^2$ with Lipschitz boundary, we study the following optimal control problem

$$\min_{(\varphi,\boldsymbol{u},p)}\int_\Omega\frac{1}{2}a\cdot\left(\mu(\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^\top)-p\,\mathrm{I}\right)\nabla\varphi+\frac{\gamma}{2c_0}\left(\frac{1}{\varepsilon}\Psi(\varphi)+\frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\mathrm{d}x$$

subject to $(\varphi,\boldsymbol{u},p)$ solving the porous-medium Navier–Stokes Eqs. 18 and the following integral constraints:

$$\begin{aligned}
\mathcal{G}_1(\varphi)&=\int_\Omega\tfrac{1}{2}(1-\varphi)x_1\,\mathrm{d}x\;=0,\\
\mathcal{G}_2(\varphi)&=\int_\Omega\tfrac{1}{2}(1-\varphi)x_2\,\mathrm{d}x\;=0,\\
\mathcal{G}_3(\varphi)&=M-\int_\Omega\tfrac{1}{2}\rho(x)(1-\varphi)\,\mathrm{d}x\;\geq 0,\\
\mathcal{G}_4(\varphi)&=\int_\Omega\varphi-\beta\,\mathrm{d}x\;\geq 0,
\end{aligned}$$

where a is a constant unit vector parallel to the flow direction $\boldsymbol{u}_\infty$, $M>0$ is a given positive constant representing an upper bound on the mass of the object,

$\rho(x) \in L^\infty(\Omega)$ is a non-negative mass density, and $\beta \in (-1, 1)$ so that the object is constraint to occupy a maximal volume of $\frac{1-\beta}{2} |\,\Omega\,|$.

The constraints $\mathcal{G}_1(\varphi) = 0$ and $\mathcal{G}_2(\varphi) = 0$ imply that the center of mass for the object is located at the origin in $\mathbb{R}^2$ (which we can assume to hold without loss of generality by translating the domain Ω). We point out that one can also consider more general surface objective functionals h that are one-homogeneous with respect to the last variable, as well as volume objective functionals b, however we consider this particular example of drag minimization as a practical application of our present approach. Furthermore, in this example we have chosen to neglect the penalization term $\hat{\alpha}_\varepsilon |\,\boldsymbol{u}\,|^2$ in the objective functional.

It is straightforward to check that the function $h(x, \nabla \boldsymbol{u}, p, \nabla\varphi) := \nabla\varphi \cdot \big(\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top) - p\,\mathrm{I}\big)\,a$ fulfills (A4) by the application of the Young's inequality. Furthermore, it is shown in [14, Proof of Thm. 4.1 and Rmk. 4.2] that the functional $\mathcal{H}(\varphi, \boldsymbol{u}, p) := \int_\Omega \nabla\varphi \cdot \big(\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top) - p\,\mathrm{I}\big)\,a\,\mathrm{dx}$ is bounded from below for $\varphi \in H^1(\Omega) \cap L^\infty(\Omega)$ with $s_a \le \varphi \le s_b$ a.e. in Ω, $\boldsymbol{u} \in \boldsymbol{H}^1_{g,\sigma}(\Omega)$ and $p \in L^2_0(\Omega)$, and satisfies due to the product of weak-strong convergence:

$$\lim_{n\to\infty} \mathcal{H}(\varphi_n, \boldsymbol{u}_n, p_n) = \mathcal{H}(\varphi, \boldsymbol{u}, p)$$

for sequences $\varphi_n \rightharpoonup \varphi$ in $H^1(\Omega)$, $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\boldsymbol{H}^1(\Omega)$ and $p_n \to p$ in $L^2(\Omega)$. Hence, (A8) is also fulfilled. A short computation shows that

$$\mathrm{D}_2 h = \mu(\nabla\varphi \otimes a + a \otimes \nabla\varphi),\ \mathrm{D}_3 h = -a \cdot \nabla\varphi,\ \mathrm{D}_4 h = (\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top) - p\,\mathrm{I})a,$$

and as a is a constant vector, one can infer that (B1) (specifically (28)) is also fulfilled. Then, it remains to verify (A5), (A6), (B2) and (C1) for the existence of Lagrange multipliers for the integral constraints $\mathcal{G}_1, \ldots, \mathcal{G}_4$, and show that the admissible set $\mathbb{K}_{ad}$ is non-empty.

For the latter, note that we have the trivial example $\phi \equiv 1 \in \mathbb{K}_{ad}$ which corresponds to the case where there is no object in the domain Ω. In the following we will construct a non-trivial example in order to rule out the possibility where $\mathbb{K}_{ad} = \{1\}$, which would imply the solution to the shape optimization problem is to have no object at all. We can always choose a function $\phi \in H^1(\Omega)$, $-1 \le \phi \le 1$ a.e. in Ω such that

$$|\,\Omega\,|\,\beta < \int_\Omega \phi\,\mathrm{dx}\,, \qquad \int_\Omega \tfrac{1}{2}(1-\phi)x_i\,\mathrm{dx} = 0 \ \text{ for } i = 1, 2,$$

which is equivalent to choosing an object $\{\phi = -1\}$ with its center of mass at the origin with volume bounded above by $\frac{1-\beta}{2} |\,\Omega\,|$. Note that the mapping $\phi \mapsto \int_\Omega \frac{1}{2}\rho(x)(1-\phi)\,\mathrm{dx}$ is continuous, and thus we can always decrease the volume of the object region $\{\phi = -1\}$ to ensure the mass is bounded from above by the constant M. This ensures that $\phi \in \mathbb{K}_{ad}$ and hence (A7) is satisfied.

As Ω is a bounded domain, the functions x_1, x_2 are bounded. Then, upon setting

$$\begin{array}{llll}
\mathcal{K}_1 = x_1, & y_1 = \frac{1}{2}(1-\varphi), & k_1 = 0, & \boldsymbol{L}_1 = \boldsymbol{0}, \\
\mathcal{K}_2 = x_2, & y_2 = \frac{1}{2}(1-\varphi), & k_2 = 0, & \boldsymbol{L}_2 = \boldsymbol{0}, \\
\mathcal{K}_3 = -\rho(x), & y_3 = \frac{1}{2}(1-\varphi), & k_3 = M\,|\,\Omega\,|^{-1}, & \boldsymbol{L}_3 = \boldsymbol{0}, \\
\mathcal{K}_4 = 1, & y_4 = \varphi, & k_4 = -\beta, & \boldsymbol{L}_4 = \boldsymbol{0},
\end{array}$$

we observe that (A5), (A6) and (B2) are fulfilled by the above choices. Then, by Thm. 1 we are guaranteed the existence of a minimizer φ_ε to the optimal control problem. To verify the main assumption (C1) and derive the optimality conditions, we have to show that for an arbitrary $z = (z_1, z_2, z_3, z_4)^\top \in \mathbb{Y} = \mathbb{R}^4$, there exists one function $\psi_* \in \Phi$, along with non-negative constants $\tau_1, \ldots, \tau_4, \xi_1, \xi_2, \eta_1, \eta_2$ such that the following four conditions are fulfilled simultaneously:

$$2z_1 = \tau_1 \int_\Omega (\varphi_\varepsilon - \psi_*) x_1 \,\mathrm{dx}\,, \quad 2z_2 = \tau_2 \int_\Omega (\varphi_\varepsilon - \psi_*) x_2 \,\mathrm{dx}\,, \tag{49a}$$

$$z_3 = \tau_3 \int_\Omega \tfrac{1}{2}\rho(x)(\psi_* - \varphi_\varepsilon)\,\mathrm{dx} - \eta_1 + \xi_1 \left(M - \int_\Omega \tfrac{1}{2}\rho(x)(1-\varphi_\varepsilon)\,\mathrm{dx} \right), \tag{49b}$$

$$z_4 = \tau_4 \int_\Omega \psi_* - \varphi_\varepsilon \,\mathrm{dx} - \eta_2 + \xi_2 \left(\int_\Omega \varphi_\varepsilon - \beta \,\mathrm{dx} \right). \tag{49c}$$

Due to their nature as equality constraints, we can use the fact that $\mathcal{G}_1(\varphi_\varepsilon) = \mathcal{G}_2(\varphi_\varepsilon) = 0$ to simplify (49a) into

$$2z_i = \tau_i \int_\Omega (1 - \psi_*) x_i \,\mathrm{dx} \text{ for } i = 1, 2. \tag{50}$$

We first argue for (50). As the origin $\mathbf{0} \notin \partial\Omega$, this implies that Ω has non-empty intersections with the four quadrants of $\mathbb{R}^2$, which we denote by $Q_1 = \{x_1, x_2 > 0\}$, $Q_2 = \{x_1 < 0, x_2 > 0\}$, $Q_3 = \{x_1, x_2 < 0\}$ and $Q_4 = \{x_1 > 0, x_2 < 0\}$. If z_1 (resp. z_2) is zero, we choose τ_1 (resp. τ_2) to be zero. Thus, it is sufficient to focus on the case where z_1 and z_2 are non-zero, and in this case we consider a function $\psi_* \in \Phi$ not identically equal to 1 with $\beta < \psi_* \leq 1$ a.e. in Ω such that the non-empty set $A := \mathrm{supp}(1 - \psi_*)$ has Lebesgue measure

$$|\, A \,| < \frac{2M}{(1-\beta)\|\rho\|_{L^\infty(\Omega)}} \tag{51}$$

and satisfies

$$A \subset\subset Q_i \cap \Omega \text{ if } (z_1, z_2) \in Q_i \text{ for } i = 1, 2, 3, 4.$$

Then, we set

$$\tau_i = \frac{2z_i}{\int_\Omega (1 - \psi_*) x_i \,\mathrm{dx}},$$

and thanks to the fact that $\psi_* \leq 1$ a.e. in Ω, the function $1 - \psi_*$ is non-negative in Ω and only positive in A. The location of A implies that the integrand $(1 - \psi_*)x_i$ has the same sign as z_i for $i = 1, 2$, and so τ_i is positive for $i = 1, 2$. The condition on the Lebesgue measure of A is used to satisfy the mass constraint.

For the inequality constraint, we have to show that the same function ψ_* considered above simultaneously satisfies (49b) and (49c). We argue for the mass constraint, and the volume constraint follows along a similar argument. There are two cases to consider: suppose the inequality constraint $\mathcal{G}_3(\varphi_\varepsilon)$ is not active for the minimizer φ_ε,

i.e., φ_ε satisfies $\int_\Omega \frac{1}{2}\rho(x)(1-\varphi_\varepsilon)\,\mathrm{dx} < M$. Then, we can choose $\tau_3 = 0$ and it holds that

$$\left\{-\eta_1 + \xi_1\left(M - \int_\Omega \tfrac{1}{2}\rho(x)(1-\varphi_\varepsilon)\,\mathrm{dx}\right) \mid \eta_1, \xi_1 \geq 0\right\} = \mathbb{R}.$$

Hence, we have fulfilled (49b) without making use of the function ψ_*. On the other hand, if $\mathcal{G}_3(\varphi_\varepsilon)$ is active, i.e., $\int_\Omega \frac{1}{2}\rho(x)(1-\varphi_\varepsilon)\,\mathrm{dx} = M$, the condition (49b) simplifies to

$$z_3 = \tau_3\left(M + \int_\Omega \tfrac{1}{2}\rho(x)(\psi_* - 1)\,\mathrm{dx}\right) - k_1.$$

A short calculation using (51) shows that the quantity in the bracket is positive, and so

$$\left\{\tau_3\left(M + \int_\Omega \tfrac{1}{2}\rho(x)(\psi_* - 1)\,\mathrm{dx}\right) - \eta_1 \mid \eta_1, \tau_1 \geq 0\right\} = \mathbb{R},$$

which implies that (49b) is fulfilled. Indeed, we see that

$$\begin{aligned} M - \int_\Omega \tfrac{1}{2}\rho(x)(1-\psi_*)\,\mathrm{dx} &= M - \int_A \tfrac{1}{2}\rho(x)(1-\psi_*)\,\mathrm{dx} \\ &\geq M - \tfrac{1}{2}\|\rho\|_{L^\infty(\Omega)}(1-\beta)\,|\,A\,| > 0. \end{aligned}$$

For the volume constraint (49c) we again divide the argument into two cases: if $\mathcal{G}_4(\varphi_\varepsilon)$ is inactive, then (49c) holds automatically without the use of the function ψ_*, and if $\mathcal{G}_4(\varphi_\varepsilon)$ is active, then using $\psi_* > \beta$ yields the desired result.

As a consequence, (4.2) is fulfilled and we obtain the existence of Lagrange multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_3, \lambda_4 \in \mathbb{R}_{\geq 0}$. By Thm. 2 the first order optimality condition is

$$\begin{aligned} 0 \leq{}& \left\langle \frac{\gamma\varepsilon}{2c_0}\nabla\varphi_\varepsilon + \frac{1}{2}\left(\mu\left(\nabla \boldsymbol{u}_\varepsilon + (\nabla \boldsymbol{u}_\varepsilon)^\top\right) - p_\varepsilon \mathrm{I}\right) a, \nabla(\zeta - \varphi_\varepsilon)\right\rangle_{L^2(\Omega)} \\ &+ \left\langle -\alpha_\varepsilon'(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + \frac{\gamma}{2c_0\varepsilon}\Psi'(\varphi_\varepsilon), \zeta - \varphi_\varepsilon\right\rangle_{L^2(\Omega)} \\ &+ \left\langle -\tfrac{1}{2}\lambda_1 x_1 - \tfrac{1}{2}\lambda_2 x_2 - \tfrac{1}{2}\lambda_3\rho(x) + \lambda_4, \zeta - \varphi_\varepsilon\right\rangle_{L^2(\Omega)} \qquad \forall \zeta \in \Phi, \end{aligned}$$

together with the complementary slackness conditions

$$\lambda_3\left(M - \int_\Omega \tfrac{1}{2}\rho(x)(1-\varphi_\varepsilon)\,\mathrm{dx}\right) = 0, \quad \lambda_4\left(\int_\Omega \varphi_\varepsilon - \beta\,\mathrm{dx}\right) = 0.$$

Remark 2 We point out that the mass constraint $\mathcal{G}_3(\varphi) = M - \int_\Omega \frac{1}{2}\rho(x)(1-\varphi)\,\mathrm{dx} \geq 0$ can also be thought of as a constraint on a construction cost, where the value $\rho(x) > 0$ represents the cost of building the object at the point $x \in \Omega$, and M denotes a maximal cost.

Let us now consider a similar model problem but with the single integral constraint on the total potential power (7). More precisely, in a bounded domain $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary, we study the following optimal control problem

$$\min_{(\varphi,\boldsymbol{u},p)} \int_\Omega \frac{1}{2} a \cdot \left(\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top) - p\,\mathrm{I}\right) \nabla\varphi + \frac{\gamma}{2c_0}\left(\frac{1}{\varepsilon}\Psi(\varphi) + \frac{\varepsilon}{2} \mid \nabla\varphi \mid^2\right) \mathrm{dx}\,,$$

subject to $(\varphi, \boldsymbol{u}, p)$ solving the porous-medium Navier–Stokes Eqs. 18 (with zero body force $\boldsymbol{f} = \boldsymbol{0}$) and the following integral constraint:

$$G(\varphi, \boldsymbol{u}) = \textstyle\int_\Omega D \mid \Omega \mid^{-1} - \frac{1}{2}(1+\varphi)\frac{\mu}{2} \mid \nabla \boldsymbol{u} \mid^2 \mathrm{dx} \geq 0,$$

where $a = -\boldsymbol{u}_\infty^\perp$ is the negative unit vector perpendicular to the flow direction $\boldsymbol{u}_\infty$ and $D > 0$ is a given positive constant representing an upper bound on the total potential power. Then, upon setting

$$\mathcal{K} = \frac{\mu}{2} \mid \nabla \boldsymbol{u} \mid^2, \quad y = -\tfrac{1}{2}(1+\varphi), \quad k = D \mid \Omega \mid^{-1}, \quad \boldsymbol{L} = \boldsymbol{0},$$

we see that (A5), (A6) and (B2) are fulfilled. In this setting we observe that the trivial example $\varphi \equiv -1$ belongs to the admissible set of design functions $\mathbb{K}_{ad}$. A non-trivial example can be found if the domain Ω is sufficiently large or the viscosity μ is sufficiently small. Indeed, let φ be a function in $H^1(\Omega)$ with $-1 \leq \varphi \leq 1$ a.e. in Ω but not identically equal to 1 or -1. Denote by $\boldsymbol{u}$ the unique velocity field associated to the state Eq. 14a, then by (22) it holds that

$$\int_\Omega \frac{1+\varphi}{2}\frac{\mu}{2} \mid \nabla \boldsymbol{u} \mid^2 \mathrm{dx} \leq \frac{\mu}{2}\|\nabla \boldsymbol{u}\|^2_{L^2(\Omega)} < \frac{\mu^3}{2K_\Omega^2}, \tag{52}$$

where from (23) the constant K_Ω is $K_\Omega = \frac{1}{2} \mid \Omega \mid^{\frac{1}{2}}$ in two dimensions. Note that the above upper bound is independent of φ. For any fixed positive constant D, we can take a sufficiently large domain Ω or sufficiently small viscosity μ, so that $\frac{\mu^3}{2K_\Omega^2} \leq D$. Then, this implies that $\mathbb{K}_{ad}$ is non-empty and thus (A7) is fulfilled. Furthermore, following the arguments above, we deduce by Thm. 1 that there exists at least one minimizer φ_ε to the optimal control problem. Writing $\mathcal{G}(\varphi_\varepsilon) = G(\varphi_\varepsilon, S_\varepsilon(\varphi_\varepsilon))$, to verify the assumption (C1), we have to show for an arbitrary $z \in \mathbb{R}$, there exists one function $\psi_* \in \Phi$ along with non-negative constants τ, ξ, η such that

$$z = \tau \mathrm{D}\mathcal{G}(\varphi_\varepsilon)(\psi_* - \varphi_\varepsilon) - \eta + \xi\mathcal{G}(\varphi_\varepsilon). \tag{53}$$

Observe that for this particular setting (with a large domain Ω or small viscosity μ), the inequality $D \geq \frac{\mu^3}{2K_{\Omega^2}}$ holds independently of the minimizer φ_ε, and thus by (52) $\mathcal{G}(\varphi_\varepsilon) > 0$ holds. In particular, the constraint is always inactive, and we do not need to find the function ψ_* as

$$\{\xi\mathcal{G}(\varphi_\varepsilon) - \eta \mid \eta, \xi \geq 0\} = \mathbb{R}.$$

Furthermore, as the constraint is always inactive the complementary slackness condition (35) implies that the associated Lagrange multiplier λ is zero. Hence, by Thm. 2 the first order optimality condition is

$$0 \leq \left\langle \frac{\gamma\varepsilon}{2c_0}\nabla\varphi_\varepsilon + \frac{1}{2}\left(\mu\left(\nabla \boldsymbol{u}_\varepsilon + (\nabla \boldsymbol{u}_\varepsilon)^\top\right) - p_\varepsilon \mathrm{I}\right) a, \nabla(\zeta - \varphi_\varepsilon)\right\rangle_{L^2(\Omega)} + \left\langle -\alpha_\varepsilon'(\varphi_\varepsilon)\boldsymbol{u}_\varepsilon \cdot \boldsymbol{q}_\varepsilon + \frac{\gamma}{2c_0\varepsilon}\Psi'(\varphi_\varepsilon), \zeta - \varphi_\varepsilon\right\rangle_{L^2(\Omega)} \quad \forall \zeta \in \Phi.$$

6 Numerical implementation and simulations

Let us now describe how we can use the above results to compute optimal shapes and topologies in given flow settings. Since our optimization variable is a phase field, and thus has the natural regularity $\varphi \in H^1(\Omega) \cap L^\infty(\Omega)$, we use the variable metric projection type (VMPT) method proposed in [5] to solve the resulting minimization problems. A standard projected gradient method can not be used for the constraint minimization problem due to the fact that $H^1(\Omega) \cap L^\infty(\Omega)$ is not a Hilbert space. The VMPT method uses derivative information which can be represented with the help of the adjoint variables as specified in (36a).

For the potential function Ψ we use the double-obstacle free energy, namely

$$\Psi(\varphi_\varepsilon) = \begin{cases} \frac{1}{2}(1-\varphi_\varepsilon^2) & \text{if } |\varphi_\varepsilon| \leq 1, \\ \infty & \text{else.} \end{cases} \tag{54}$$

From this we obtain the constraint $|\varphi_\varepsilon| \leq 1$, and $c_0 = \frac{\pi}{2}$, where c_0 is the constant defined in (13). Although the double-obstacle potential (54) does not satisfy (A1), the analysis is not affected once we choose $s_a = -1$ and $s_b = 1$, so that $|\varphi_\varepsilon| \leq 1$ and the potential becomes $\Psi(\varphi_\varepsilon) = \frac{1}{2}(1-\varphi_\varepsilon^2)$. We refer the reader to [11, 13] which also uses the double-obstacle potential (54). For the porous-medium term $\alpha_\varepsilon(\varphi_\varepsilon)$ in the state Eqs. 14a we choose

$$\alpha_\varepsilon(\varphi_\varepsilon) = \frac{\overline{\alpha}}{2\varepsilon}(1-\varphi_\varepsilon), \tag{55}$$

with a fixed positive constant $\overline{\alpha}$, and (A0) is fulfilled with $s_a = -1$ and $s_b = 1$. We choose

$$\hat{\alpha}_\varepsilon \equiv \alpha_\varepsilon$$

so that (A2) is also satisfied. For the remaining part of this section, we denote both variables by α_ε, set $\boldsymbol{f} = \boldsymbol{0}$ in (14a), and define

$$\Phi = \{f \in H^1(\Omega) \mid -1 \leq f \leq 1 \text{ a.e. in } \Omega\}.$$

6.1 Spatial discretization

We use finite elements for the numerical discretization of the minimization problem. We use piecewise linear and globally continuous finite elements for the

representation of φ_ε, p_ε and π_ε and piecewise quadratic and globally continuous finite elements for $\boldsymbol{u}_\varepsilon$ and q_ε on a conforming triangulation of the domain Ω.

It is well-known that in phase field applications the variable φ_ε changes rapidly across the interfacial layers, and an adaptive concept for its spatial resolution is indispensable. Hence, for the mesh generation we use the Dual Weighted Residual (DWR) method [2] where our implementation is guided by [20]. This generates adaptive meshes which well resolve the interfacial regions, and also well reflect the underlying flow physics, compare also [19]. The DWR approach is only applicable if for a given triangulation an optimal solution is already found and uses this information to calculate error indicators.

For fast calculations, it is desirable to use coarse meshes. In the core of the VMPT method we solve projection-type problems using a primal-dual-active-set strategy (PDAS). Here the active set corresponds to degrees of freedom with $|\varphi_\varepsilon| = 1$. Thus in every step of the PDAS we solve the problems on the inactive set $|\varphi_\varepsilon| < 1$ only. Note that the integral constraints have to be fulfilled by changing the phase field on the inactive set only. If this set contains too few degrees of freedom, the PDAS is not successful in solving the projection-type problem and thus the algorithm breaks down.

To overcome this numerical issue on coarse meshes, we additionally require that a given amount, say 2%, of the phase field's degrees of freedom are inactive. If this is not the case, we use mesh adaptation that is based on φ_ε only, namely we use the jumps of the normal derivatives of φ_ε across edges as proposed in [13] to generate new degrees of freedom inside the interface to be able to proceed with the PDAS.

We stop the adaptation loop as soon as a given maximum number of degrees of freedom is reached.

6.2 Topology optimization - a tube through heavy ground

Although we have mainly focused on shape optimization with the phase field approach in this paper, we point out that using a phase field variable for the representation of the unknown shape also allows us to deal with situations where no a priori toplogical information is available. In particular, the phase field approach is capable of topology optimization, as done in [4, 27, 35, 38]. Here we consider the situation where the domain $\Omega = (0,1)^2$ contains several impermeable rocks and we would like to search for a tube that connects the inflow at the bottom to the outflow at the top, see Fig. 1.

Constructing a tube through the rocks is expensive and therefore a tube that avoids these regions is desired. So this is a setting where we want to minimize the cost of an object. The inflow and the outflow regions as well as the location of the rocks are a priori known. We define the inflow and outflow conditions as

$$g_{in}(x) = \begin{pmatrix} 0 \\ \max\left(2\left(1-\left(\frac{x_1-0.5}{1/6}\right)^2\right), 0.0\right) \end{pmatrix},$$

$$g^i_{out}(x) = \begin{pmatrix} \max\left((-1)^i\left(1-\left(\frac{x_2-0.8}{1/12}\right)^2\right), 0.0\right) \\ 0 \end{pmatrix} \text{ for } i = 1, 2.$$

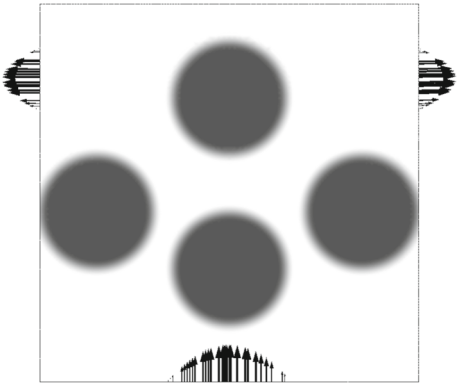

Fig. 1 The inflow and outflow conditions for the Navier–Stokes equations together with the location of the rocks

For the objective functional we define a 'rock' centered at $\boldsymbol{m}$ with radius σ and associated cost c as

$$R[\boldsymbol{m}, \sigma, c](x) := (c-1)\left(\frac{\phi_0(-\frac{1}{\varepsilon}(\|\frac{x-\boldsymbol{m}}{\sigma}\|-1))+1}{2}\right)+1,$$

$$\text{where} \quad \phi_0(z) = \begin{cases} \sin(z) & \text{if } |z| \leq \frac{\pi}{2}, \\ \operatorname{sign}(z) & \text{else.} \end{cases}$$

We consider the functions

$$b(x, \boldsymbol{u}, \nabla \boldsymbol{u}, p, \varphi) := \left(\tfrac{1+\varphi}{2}\right) \prod_{i=1}^{4} R[\boldsymbol{m}_i, \sigma, c](x), \quad h(x, \nabla \boldsymbol{u}, p, \nabla \varphi) = 0,$$

where

$$\begin{aligned} \boldsymbol{m}_1 &= (0.5, 0.3)^\top, & \boldsymbol{m}_2 &= (0.15, 0.45)^\top, \\ \boldsymbol{m}_3 &= (0.85, 0.45)^\top, & \boldsymbol{m}_4 &= (0.5, 0.75)^\top. \end{aligned}$$

The optimization problem (17) then becomes

$$\begin{aligned} \min_{(\varphi, \boldsymbol{u}, p)} \mathcal{J}_\varepsilon(\varphi, \boldsymbol{u}, p) = & \int_\Omega \frac{1}{2}\alpha_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2 + \frac{1+\varphi}{2}\prod_{i=1}^{4} R[\boldsymbol{m}_i, \sigma, c]\,\mathrm{dx} \\ & + \int_\Omega \frac{\gamma}{\pi}\left(\frac{1}{\varepsilon}\Psi(\varphi) + \frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\mathrm{dx}\,, \end{aligned}$$

subject to $\varphi \in \Phi$, $\boldsymbol{u} \in H^1_{g,\sigma}(\Omega)$, $p \in L^2_0(\Omega)$ satisfying (18), and α_ε was defined earlier in (55). For this example we do not apply any integral constraints, as it serves to demonstrate the strength of the phase field approach in being able to deal with situations where no a priori topological information is known. Having the solution to the unconstrained problem at hand, one might reduce for example the size of the tube by imposing additional volume constraints. However, specifying such constraints beforehand might lead to inadmissible situations.

We start the optimization procedure with no prior information, i.e., $\varphi_\varepsilon^0 \equiv 0$, on a homogeneous coarse grid with mesh size $h = 1/20$ yielding 685 degrees of unknowns for φ_ε. We stop the solution and adaptation procedures as soon as an optimal solution with more than 100000 degrees of freedom is found. The numerical

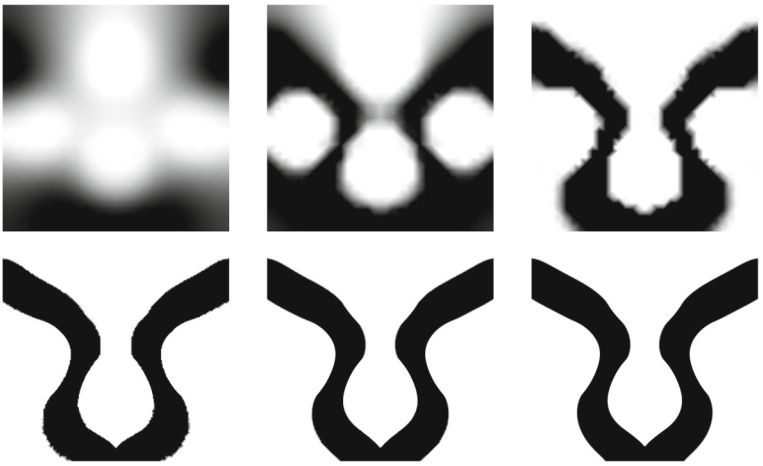

Fig. 2 The iterations 20,30,40,80,120,190 of the VMPT to minimize the cost of a tube through heavy ground. We see that after 40 steps already an optimal structure is found and that in subsequent steps mostly the resolution of the structure is improved. Let us also note that at iteration 20 we have only 923 degrees of freedom for φ_ε and in iteration 40 still only 1398 degrees of freedom. The final iteration has 125069 degrees of freedom

parameters are: $\sigma = 0.15$, $c = 50$, $\varepsilon = 0.01$, $\overline{\alpha} = 5$, $\mu = 0.02$ and $\gamma = 0.001$. To stress the advantages of our approach in Fig. 2 we show φ_ε at various stages of the optimization procedure.

6.3 Reproduction of results on drag minimization from earlier works

We now reproduce the numerical results for the surface formulation of drag minimization presented by the authors in [14]. The key distinction is that in [14] a gradient flow approach is used to solve the optimality conditions, leading to a non-linear time-dependent equation of Cahn–Hilliard type for φ_ε. However, here we employ the VMPT method to solve the optimization problem, which reads

$$\begin{aligned}\min_{(\varphi,\boldsymbol{u},p)} \mathcal{J}_\varepsilon(\varphi,\boldsymbol{u},p) = &\int_\Omega \tfrac{1}{2}\alpha_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2 + \frac{\gamma}{\pi}\left(\frac{1}{\varepsilon}\Psi(\varphi) + \frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\mathrm{dx}\\ &+ \int_\Omega \tfrac{1}{2}a\cdot\left(\mu\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top\right) - p\,\mathrm{I}\right)\nabla\varphi\,\mathrm{dx}\end{aligned}$$

subject to $\varphi \in \Phi$, $\boldsymbol{u} \in H^1_{g,\sigma}(\Omega)$, $p \in L^2_0(\Omega)$ satisfying (18) and the volume constraint (see (9))

$$\int_\Omega \varphi\,\mathrm{dx} \le \beta_2\,|\,\Omega\,| \text{ for } \beta_2 \in (-1,1). \tag{56}$$

We use the parameters from [14], namely $\Omega = (0,1.7)\times(0,0.4)$, $\varepsilon = 0.00025$, $\overline{\alpha} = 0.03$, $\mu = 0.001$ and $\gamma = 0.01$. The boundary velocity is set to $g = (1,0)^\top$ to stay close to the analysis and we initialize the optimization with $\varphi^0_\varepsilon(x) :=$

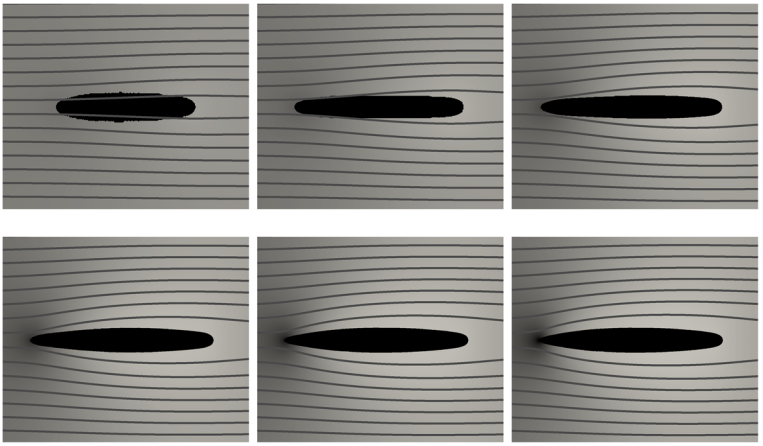

Fig. 3 The optimal shapes for the minimization of drag in the surface formulation with the parameters from [14] for $\varepsilon = 0.008, 0.004, 0.002, 0.001, 0.0005, 0.00025$ (left upper to right lower). Inflow from the left with $g \equiv (1, 0)^\top$. The shape is shown in black. The pressure is shown in gray, where darker gray means larger pressure, and some streamlines of the velocity are shown in black

$-R[(0.5, 0.2)^\top, 0.25, -1](x)$, i.e., a ball around $m = (0.5, 0.2)^\top$ with radius $r = 0.25$. For the volume constraint, we choose $\beta_2 = \beta = 0.975$, i.e., $\int_\Omega \varphi \, \mathrm{dx} \leq 0.663$.

To be able to use only a small number of unknown as long as possible, we start the optimization with $\varepsilon = 0.008$ and a maximum number of allowed degrees of freedom of 10000. We halve the value of ε as soon as an optimal solution is found with current maximum allowed number of degrees of freedom and increase this value by 20%, resulting in 45000 unknowns for the final result. In Fig. 3 we show the optimal shape for different values of ε, namely $\varepsilon \in \{0.008, 0.004, 0.002, 0.001, 0.0005, 0.00025\}$. In Table 1 we show the diffuse interface drag

$$F_\varepsilon^D = \int_\Omega \tfrac{1}{2} a \cdot \left(\mu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top \right) - p \, \mathrm{I} \right) \nabla \varphi \, \mathrm{dx} \tag{57}$$

and the sharp interface drag

$$F^D = \int_\Gamma a \cdot \left(\mu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top \right) - p \, \mathrm{I} \right) \nu \, \mathrm{d}\mathcal{H}^{d-1} \tag{58}$$

by evaluation with $a = (1.0, 0.0)^\top$ over $\Gamma = \{\varphi_\varepsilon = 0\}$.

We reproduce the results found in [14] where a gradient flow approach is applied that is based on an artificial time evolution. We stress that, in using a gradient flow approach, the interface has to be resolved in each time step of the temporal evolution, which leads to a large numerical effort. To be precise, while the results shown here are found in a few hours using the VMPT method, the results in [14] required several days of calculation using the gradient flow.

Table 1 The diffuse (F_ε^D) and sharp (F^D) drag for the parameters from [14] and different values of ε. Note that $\alpha_\varepsilon(-1) \to \infty$ for $\varepsilon \to 0$, i.e., the object becomes less permeable and thus the drag increases with $\varepsilon \to 0$. In [14] for $\varepsilon = 0.00025$ we observed $F_\varepsilon^D = 3.9117 \times 10^{-2}$ and $F^D = 3.9499 \times 10^{-2}$

ε	0.008	0.004	0.002
F_ε^D	1.0570×10^{-2}	1.9806×10^{-2}	2.8370×10^{-2}
F^D	1.1103×10^{-2}	2.0519×10^{-2}	2.9025×10^{-2}
ε	0.001	0.0005	0.00025
F_ε^D	3.4255×10^{-2}	3.8184×10^{-2}	4.0739×10^{-2}
F^D	3.4777×10^{-2}	3.8572×10^{-2}	4.1012×10^{-2}

6.4 Comparison of volume and surface formulations for drag

Let us point out that the hydrodynamic force component (58) in its classical representation as a surface integral over Γ can be expressed in terms of a volume integral over the fluid region E, and this reformulation has been used extensively in numerical simulations, see [9, §5.1], [15, §2.2], and [22, §9]. Given the unit vector $a \neq \mathbf{0}$, let η be a smooth vector field such that

$$\eta = a \text{ on } \Gamma \text{ and } \eta = \mathbf{0} \text{ on } \partial\Omega. \tag{59}$$

This can be done since it is assumed in the introduction that Γ does not intersect with $\partial\Omega$. Then, by taking the scalar product of (4a) with η, we obtain

$$0 = \int_E -\operatorname{div}\,(\mu\nabla \boldsymbol{u}) \cdot \eta + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\eta + \nabla p \cdot \eta - \boldsymbol{f}\cdot\eta \,\mathrm{dx}\,.$$

Integrating by parts and noting that the boundary integrals over $\partial\Omega$ vanish owning to $\eta = \mathbf{0}$ on $\partial\Omega$ yields

$$\begin{aligned} &\textstyle\int_\Gamma a \cdot \left(\mu\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top\right) - p\,\mathrm{I}\right)\nu \,\mathrm{d}\mathcal{H}^{d-1} = \int_\Gamma a\cdot(\mu\nabla\boldsymbol{u} - p\,\mathrm{I})\,\nu \\ &\quad = -\textstyle\int_E \mu\nabla\boldsymbol{u}\cdot\nabla\eta + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\eta - p\operatorname{div}\eta - \boldsymbol{f}\cdot\eta\,\mathrm{dx}\,. \end{aligned} \tag{60}$$

Here we have also used that $\boldsymbol{u}$ has no tangential component on Γ due to the no-slip condition $\boldsymbol{u} = \mathbf{0}$ on Γ, and together with the divergence-free condition, we obtain that $(\nabla\boldsymbol{u})^\top\nu = \mathbf{0}$ on Γ (see [14, §2] for more details). This implies that we can also consider the following function as the volume formulation of the drag (if a is parallel to the flow direction)

$$\int_\Omega -\tfrac{1}{2}(1+\varphi)\,(\mu\nabla\boldsymbol{u}\cdot\nabla\eta + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\eta - p\operatorname{div}\eta - \boldsymbol{f}\cdot\eta)\,\mathrm{dx}\,. \tag{61}$$

Alternatively, using integration by parts and the boundary conditions $\eta = \mathbf{0}$ on $\partial\Omega$ and $\boldsymbol{u} = \mathbf{0}$ on Γ, we see that

$$\int_E (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\cdot\eta\,\mathrm{dx} = -\int_E (\boldsymbol{u}\cdot\nabla)\eta\cdot\boldsymbol{u}\,\mathrm{dx}\,,$$

and so we may also use the function

$$\int_\Omega -\tfrac{1}{2}(1+\varphi)\,(\mu\nabla \boldsymbol{u}\cdot\nabla\eta-(\boldsymbol{u}\cdot\nabla)\eta\cdot\boldsymbol{u}-p\,\mathrm{div}\,\eta-\boldsymbol{f}\cdot\eta)\,\mathrm{dx}\,, \tag{62}$$

as a volume formulation for the drag. The corresponding phase field approximations of (61) and (62) have the exact same form. However, for our numerical investigations, we use the formation (62) instead of (61).

The aim of this section is to compare results of Section 6.3 with the following drag minimization problem

$$\begin{aligned}\min_{(\varphi,\boldsymbol{u},p)} \mathcal{J}_\varepsilon(\varphi,\boldsymbol{u},p) = &\int_\Omega \frac{1}{2}\alpha_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2+\frac{\gamma}{\pi}\left(\frac{1}{\varepsilon}\Psi(\varphi)+\frac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\mathrm{dx}\\ &+\int_\Omega -\tfrac{1}{2}(1+\varphi)\,(\mu\nabla\boldsymbol{u}\cdot\nabla\eta-(\boldsymbol{u}\cdot\nabla)\eta\cdot\boldsymbol{u}-p\,\mathrm{div}\,\eta-\boldsymbol{f}\cdot\eta)\,\mathrm{dx}\end{aligned}$$

subject to $\varphi\in\Phi$, $\boldsymbol{u}\in\boldsymbol{H}^1_{g,\sigma}(\Omega)$, $p\in L^2_0(\Omega)$ satisfying (18) and the volume constraint (56). In particular, we compare the optimal shapes obtained with volume formulation (62) and those obtained with the surface formulation (57) for the drag. For the above optimization problem, we consider the same setup as in Section 6.3 and set $\eta\equiv a$ on $(0.15,1.0)\times(0.13,0.27)$.

Using the surface formulation (57) we observe that for larger values of $\overline{\alpha}$ (the constant in (55)) an interfacial region $\{|\,\varphi_\varepsilon\,|<1\}$ that is neither fluid nor object appearing in front of the object. A similar behavior was observed in the previous work [14] with another minimization algorithm. In any case, a sufficiently impermeable object can be obtained by using smaller values of ε. We stress that, in Section 6.3 for $\varepsilon=0.00025$ the the velocity $|\,\boldsymbol{u}_\varepsilon\,|$ inside the object is five orders of magnitude smaller than outside the object (see [14, Fig. 1]).

On the other hand, using the volume formulation (62) we have to define the extension of the unit vector field a, namely the vector field η which has to vanish at $\partial\Omega$. We define η as the solution of a Poisson problem on Ω with $\eta=a$ on a square around the object and $\eta=0$ on $\partial\Omega$. That is, let S denote a square such that $\{\varphi_\varepsilon=-1\}\subset S$ and $\partial S\cap\partial\Omega=\emptyset$, then we solve

$$-\Delta\eta=\mathbf{0}\text{ in }\Omega\setminus S,\quad \eta=\mathbf{0}\text{ on }\partial\Omega,\quad \eta=a\text{ in }\overline{S}. \tag{63}$$

For small values of $\overline{\alpha}$, we observe that the object splits and the solid is collected close to the inflow/outflow boundaries. We believe this behavior is due to the following: On the one hand, due to the boundary condition $\eta=\mathbf{0}$ on $\partial\Omega$, the magnitude $|\,\eta\,|$ is small close to the inflow and outflow boundaries, which results in small drag forces. On the other hand, for $\overline{\alpha}$ small, the porous-medium penalization term $\int_\Omega\alpha_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2\,\mathrm{dx}$ is small, and thus the value of the objective functional can be reduced by placing material in regions where $|\,\eta\,|$ is small. Therefore, in contrast to the surface formulation (57), large values of $\overline{\alpha}$ are needed for the volume formulation to obtain reasonable optimal shapes, which additionally allows us to construct sufficiently impermeable objects when we use larger values of ε.

We use the setup from Section 6.3 with only one modification, that we set $\mu=0.01$. In Fig. 4 the optimal shapes of the objects using the surface and the volume formulations of the drag are shown. We observe that the front of the object

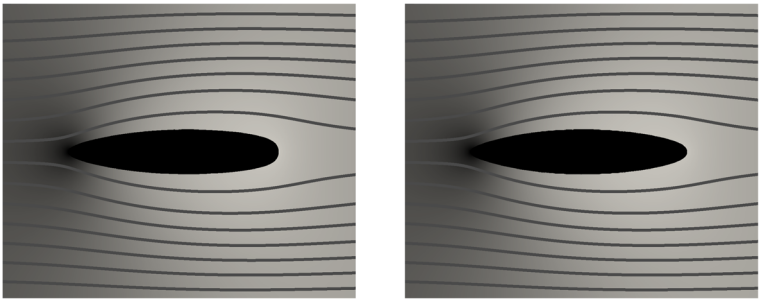

Fig. 4 The optimized shapes of the object using the surface formulation ((57), left) and the volume formulation ((62), right) of the drag with $\mu = 0.01$ and $\overline{\alpha} = 0.03$. We observe that the rear of the object is slightly more pronounced when the volume formulation is used, while the drag measured on the zero level-line in both cases is nearly identical

with both formulations is rather similar, while the surface formulation leads to a less pronounced rear. The corresponding drag values in sharp interface evaluation (58) as defined in Section 6.3 are $F^D = 0.106467052$ (volume formulation) and $F^D = 0.106470276$ (surface formulation).

As described above, using the volume formulation we can use larger values for $\overline{\alpha}$ to model objects with smaller permeability. To show the influence of $\overline{\alpha}$ in Fig. 5 we show the optimal shape for the above parameters, but using a larger value $\overline{\alpha} = 1$ and $\mu = 0.01$ (left) and $\mu = 0.001$ (right). For $\mu = 0.01$ we observe, that we get a sharper rear of the object, while the magnitude of the velocity inside the object is of order 10^{-4}, which is two orders of magnitudes smaller than in the case $\overline{\alpha} = 0.03$. We also mention that the shapes obtained here bear similarities to the optimized shape for the minimization of the dissipative energy, as presented in [13, Figs. 4 and 5]. For $\overline{\alpha} = 1$, and $\mu = 0.001$ we observe a symmetric airfoil shape.

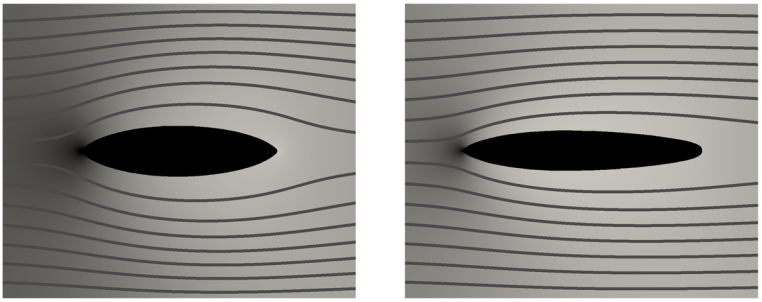

Fig. 5 The optimized shape of the object using the volume formulation and $\overline{\alpha} = 1$ with $\mu = 0.01$ (left) and $\mu = 0.001$ (right). Compared to Fig. 4 we observe a sharper rear and for $\mu = 0.001$ a symmetric airfoil shape emerges. For $\mu = 0.01$ the drag is $F^D = 0.205542595$ and the velocity inside the object is of order 10^{-4}, which is two orders of magnitude smaller than in the case $\overline{\alpha} = 0.03$. For $\mu = 0.001$ the drag is $F^D = 0.041090517$ and the velocity inside the object is of order 10^{-6}

6.5 Maximizing the lift with constraints on the total potential power

We give an example of dealing with a state constraint, namely we consider the maximization of the lift of an object under the constraint that the total potential power is bounded by some given value. This is a non-linear constraint on the state variables of the constraint optimization problem, namely the velocity field. To treat the highly non-linear potential power constraint we use Moreau–Yosida relaxation.

The optimization problem we solve is

$$\begin{aligned}\min_{(\varphi,\boldsymbol{u},p)} \mathcal{J}_\varepsilon^s(\varphi) := & \int_\Omega \tfrac{1}{2}\alpha_\varepsilon(\varphi)\,|\,\boldsymbol{u}\,|^2 + \tfrac{a}{2}\cdot\left(\mu\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top - p\,\mathrm{I}\right)\nabla\varphi\right)\,\mathrm{dx} \\ & + \int_\Omega \tfrac{\gamma}{\pi}\left(\tfrac{1}{\varepsilon}\Psi(\varphi) + \tfrac{\varepsilon}{2}\,|\,\nabla\varphi\,|^2\right)\,\mathrm{dx} \\ & + \tfrac{s}{2}\max\left(0.0,\ \int_\Omega \tfrac{1+\varphi}{2}\tfrac{\mu}{2}\,|\,\nabla\boldsymbol{u}\,|^2 - D\,|\,\Omega\,|^{-1}\,\mathrm{dx}\right)^2,\end{aligned}$$

subject to $\varphi \in \Phi$, $\boldsymbol{u} \in H^1_{g,\sigma}(\Omega)$, $p \in L^2_0(\Omega)$ satisfying (18), where to realize the maximization of lift, we set $a = (-1, 0)^\top$ as the negative unit vector perpendicular to the flow direction, and the parameter $s > 0$ penalizes violation of the constraint that the total potential power of the fluid region must be less than or equal to a prescribed value D:

$$\int_\Omega \frac{1+\varphi}{2}\frac{\mu}{2}\,|\,\nabla\boldsymbol{u}\,|^2\,\mathrm{dx} \le D.$$

The integral constraint for this optimization problem are volume constraints of the form $0.663 \le \int_\Omega \varphi_\varepsilon\,\mathrm{dx} \le 0.665$ and the center of mass is fixed at $(0.5, 0.2)^\top$.

The set up is similar to that in Section 6.3, where we set $\Omega = (0.0, 1.7)\times(0.0, 0.4)$, $g = (1, 0)^\top$, $\varphi_\varepsilon^0(x) := -R[(0.5, 0.2)^\top, 0.25, -1](x)$, i.e., a circle around $\boldsymbol{m} = (0.5, 0.2)^\top$ with radius $r = 0.25$. For the penalization parameter we choose $s = 100$ and further numerical parameters are $\varepsilon = 0.02$, $\overline{\alpha} = 2$, $\mu = 0.01$, $\gamma = 0.001$, and $D = 0.06$.

In Fig. 6 we show the resulting optimal shape of the object. As expected we observe an inclined structure in order to maximize lift, but due to the constraint on the potential power, the angle of attack is restricted. This is consistent with previous results in [14, Fig. 2].

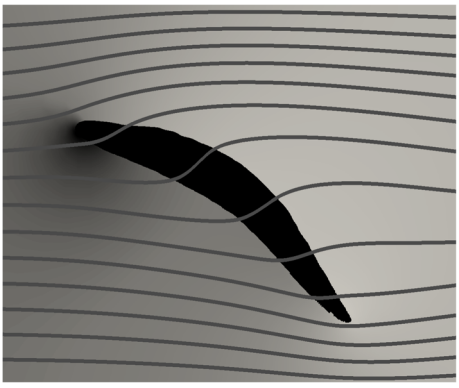

Fig. 6 The optimal shape for the maximization of the lift of an object, under a constraint on the dissipative power. We observe an inclined shape

7 Conclusion

In this paper, we formulate and analyze a phase field approximation for an abstract shape optimization problem subject to stationary Navier–Stokes flow with general objective functionals and integral state constraints. We provide examples for the objective functionals and integral constraints that are of practical relevance, and we establish the existence of minimizers, and derive the first order optimality conditions. A crucial point in the analysis is to show the existence of Lagrange multipliers corresponding to the integral constraints. In the general setting we assume that the Zowe–Kurcyusz constraint qualification holds, and verify these assumptions for two specific examples. The first involves integral constraints only in the variable φ, while the second involves the state variable $\boldsymbol{u}$. The optimality conditions are solved using the VMPT method and several simulations are performed. We demonstrate that the proposed phase field approach can handle topology optimization, and compare the results of drag minimization with previous works. Lastly, we consider an example with an integral constraint involving the state variables, namely maximization of lift with constraint on the potential power. The optimal shapes obtained are consistent with previous works on the lift-to-drag ratio for fluid flow with small Reynolds number.

Acknowledgments The authors gratefully acknowledge the financial support by the Deutsche Forschungsgemeinschaft (DFG) through the grants GA695/6-2 (first and fourth author) and HI689/7-1 (second and third author) within the priority program SPP1506 "Transport processes at fluidic interfaces". The third author additionally gratefully acknowledges the support by the DFG through the International Research Training Group IGDK 1754 "Optimization and Numerical Analysis for Partial Differential Equations with Nonsmooth Structures".

References

1. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. Oxford University Press, USA (2000)
2. Becker, R., Rannacher, R.: An optimal control approach to a posteriori error estimation in finite element methods. Acta Numer. **10**, 1–102 (2001)
3. Bello, J.A., Fernández-Cara, E., Lemoine, J., Simon, J.: The differentiability of the drag with respect to the variations of a Lipschitz domain in a Navier–Stokes flow. SIAM J. Control Optim. **35**(2), 626–640 (1997)
4. Blank, L., Hecht, C., Garcke, H., Rupprecht, C.: Sharp interface limit for a phase field model in structural optimization. SIAM J. Control Optim. **54**, 1558–1584 (2016)
5. Blank, L., Rupprecht, C.: An extension of the projected gradient method to a Banach space setting with application in structural topology optimization. SIAM J. Control Optim. **55**, 1481–1499 (2017)
6. Boisgérault, S., Zolésio, J.: Shape derivative of sharp functionals governed by Navier–Stokes flow. In: Jäger, W., Nečas, J., John, O., Najzar, K., Stará, J. (eds.) Partial Differential Equations: Theory and Numerical Solution, pp. 49–63. Chapman and Hall/CRC (1993)
7. Borrvall, T., Petersson, J.: Topology optimization of fluids in Stokes flow. Internat. J. Numer. Methods Fluids **41**(1), 77–107 (2003)
8. Bourdin, B., Chambolle, A.: Design-dependent loads in topology optimization. ESAIM Control Optim. Calc. Var. **9**, 19–48 (2003)
9. Brandenburg, C., Lindemann, F., Ulbrich, M., Ulbrich, S.: A Continuous Adjoint Approach to Shape Optimization for Navier Stokes Flow. In: Kunisch, K., Sprekels, J., Leugering, G., Tröltzsch,

F. (eds.) Optimal Control of Coupled Systems of Partial Differential Equations, International Series of Numerical Mathematics, vol. 158, pp. 35–56. Birkhäuser (2009)
10. Evans, L., Gariepy, R.: Measure theory and fine properties of functions. Studies in advanced mathematics. CRC Press, Boca Raton (1992)
11. Garcke, H., Hecht, C.: Applying a phase field approach for shape optimization of a stationary Navier-Stokes flow. ESAIM: Control Optim. Calc. Var. **22**, 309–337 (2016)
12. Garcke, H., Hecht, C.: Shape and topology optimization in Stokes flow with a phase field approach. Appl. Math. Optim. **73**, 23–70 (2016)
13. Garcke, H., Hecht, C., Hinze, M., Kahle, C.: Numerical approximation of phase field based shape and topology optimization for fluids. SIAM J. Sci. Comput. **37**(4), A1846–A1871 (2015)
14. Garcke, H., Hecht, C., Hinze, M., Kahle, C., Lam, K.F.: Shape optimization for surface functionals in Navier–Stokes flow using a phase field approach. Interfaces Free Bound. **18**(2), 219–261 (2016)
15. Giles, M., Larson, M., Levenstam, M., Süli, E.: Adaptive error control for finite element approximations of the lift and drag coefficients in viscous flow. Technical Report NA-79/06 Oxford University Computing Laboratory (1997)
16. Giusti, E.: Minimal surfaces and functions of bounded variation, Monographs in mathematics, vol. 80. Birkhäuser, Basel (1984)
17. Goldberg, H., Kampowsky, W., Tröltzsch, F.: On Nemytskij operators in L_p-spaces of abstract functions. Math. Nachr. **155**, 127–140 (1992)
18. Hecht, C.: Shape and topology optimization in fluids using a phase field approach and an application in structural optimization. Ph.D. thesis, University of Regensburg (2014)
19. Hintermüller, M., Hinze, M., Kahle, C.: An adaptive finite element Moreau–Yosida-based solver for a coupled Cahn–Hilliard/Navier–Stokes system. J. Comput. Phys. **235**, 810–827 (2013)
20. Hintermüller, M., Hinze, M., Kahle, C., Keil, T.: A goal-oriented dual-weighted adaptive finite elements approach for the optimal control of a Cahn–Hilliard–Navier–Stokes system. Preprint Hamburger Beiträge zur Angewandten Mathematik 2016-25 (2016)
21. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints. Mathematical Modelling: Theory and Applications. Springer, Netherlands (2009)
22. Hoffman, J., Johnson, C.: Adaptive finite element methods for incompressible fluid flow. In: Barth, T., Deconinck, H. (eds.) Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics, vol. 25, pp. 95–157. Springer, Berlin Heidelberg (2003)
23. Kawohl, B., Pironneay, O., Tartar, L., Zolesio, J.P.: Optimal Shape Design: Lectures Given at the Joint C.I.M./C.I.M.E. Summer School Held in Troia (Portugal), June 1-6, 1998. Lecture Notes in Mathematics / C.I.M.E. Foundation Subseries. Springer-Verlag, Berlin Heidelberg (2000)
24. Kondoh, T., Matsumori, T., Kawamoto, A.: Drag minimization and lift maximization in laminar flows via topology optimization employing simple objective function expressions based on body force integration. Struct. Multidiscip. Optim. **45**(5), 693–701 (2012)
25. Modica, L.: The gradient theory of phase transitions and the minimal interface criterion. Arch. Ration. Mech. Anal. **98**(2), 123–142 (1987)
26. Murat, F.: Contre-exemples pour divers problèmes où le contrôle intervient dans les coefficients. Ann. Mat. Pura Appl., Serie 4 **112**(1), 49–68 (1977)
27. Penzler, P., Rumpf, M., Wirth, B.: A phase-field model for compliance shape optimization in nonlinear elasticity. ESAIM: Control Optim. Calc. Var. **18**, 229–258 (2012)
28. Pironneau, O.: On optimum design in fluid mechanics. J. Fluid Mech. **64**, 97–110 (1974)
29. Plotnikov, P., Sokolowski, J.: Shape derivative of drag functional. SIAM J. Control Optim. **48**(7), 4680–4706 (2010)
30. Robinson, S.: Stability theorems for systems of inequalities, Part II: Differentiable nonlinear systems. SIAM J. Numer. Anal. **13**(4), 497–513 (1976)
31. Schmidt, S., Schulz, V.: Shape derivatives for general objective functions and the incompressible Navier–Stokes equations. Control Cybernet. **39**(3), 677–713 (2010)
32. Simon, J.: Domain variation for drag in Stokes flow. In: Control Theory of Distributed Parameter Systems and Applications, Lecture Notes in Control and Information Sciences, vol. 159. Springer, Berlin (1991)
33. Sohr, H.: The Navier–Stokes Equations: An Elementary Functional Analytic Approach. Birkhäuser Advanced Texts. Springer-Verlag, Berlin (2001)
34. Sturm, K., Hintermüller, M., Hömberg, D.: Distortion compensation as a shape optimization problem for a sharp interface model. Comput. Optim. Appl. **64**, 557–588 (2016)

35. Takezawa, A., Nishiwaki, S., Kitamura, M.: Shape and topology optimization based on the phase field method and sensitivity analysis. J. Comput. Phys. **229**, 2697–2718 (2010)
36. Tartar, L.: Problemes de Controle des Coefficients Dans des Equations aux Derivees Partielles. In: Bensoussan, A., Lions, J. (eds.) Control Theory, Numerical Methods and Computer Systems Modelling, Lecture Notes in Economics and Mathematical Systems, vol. 107, pp. 420–426. Springer, Berlin Heidelberg (1975)
37. Tröltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods, and Applications. Graduate studies in mathematics. AMS, Providence (2010)
38. Wang, M., Zhou, S.: Multimaterial structural topology optimization with a generalized Cahn–Hilliard model of multiphase transition. Struct. Multidisc. Optim. **33**, 89–111 (2007)
39. Zowe, J., Kurcyusz, S.: Regularity and stability for the mathematical programming problem in banach spaces. Appl. Math. Optim. **5**, 49–62 (1979)