

Numerical analysis of a topology optimization problem for Stokes flow

I.P.A. Papadopoulos^{*,1}, E. Süli

Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK

ARTICLE INFO

Article history:

Received 8 February 2021

Received in revised form 17 February 2022

Keywords:

Topology optimization

Stokes flow

Regularity

Finite element method

Nonconvex variational problem

Multiple solutions

ABSTRACT

Borrvall and Petersson (2003) developed the first model for the topology optimization of fluids in Stokes flow. They proved the existence of minimizers in the infinite-dimensional setting and showed that a suitably chosen finite element method will converge in a weak(-*) sense to an unspecified solution. In this work, we prove novel regularity results and extend their numerical analysis. In particular, given an isolated local minimizer to the infinite-dimensional problem, we show that there exists a sequence of finite element solutions, satisfying necessary first-order optimality conditions, that *strongly* converges to it. We also provide the first numerical investigation into convergence rates.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Topology optimization has become an effective technique in structural and additive manufacturing and has found multiple uses in medicine, architecture, and engineering [1–3]. The objective is to find the optimal distribution of a fluid or solid within a given domain that minimizes a problem-specific cost functional [4,5]. In this paper we consider a model for topology optimization for fluids proposed by Borrvall and Petersson [6]. Their seminal work has become the foundation for a number of developments in recent years [7–18]. Their goal was to minimize the power dissipation of a fluid that satisfies both the Stokes equations and a volume constraint restricting the proportion of the domain that the fluid can occupy. In their paper, they derived *generalized Stokes equations*, which incorporate the classical velocity and pressure terms but also introduce a variable, ρ , that represents the material distribution of the fluid over the given domain. The presence of fluid is indicated by a value of one in the material distribution whereas absence of fluid is represented by a value of zero. It would be ideal for $\rho : \Omega \rightarrow \{0, 1\}$, in order to remove any ambiguity in the solutions, however, in general, this is a numerically intractable objective. In the Borrvall–Petersson model $\rho : \Omega \rightarrow [0, 1]$, but the model is regularized with an inverse permeability term, α , which favors solutions where ρ is close to zero or one. From the generalized Stokes equations, Borrvall and Petersson formulated an infinite-dimensional nonconvex optimization problem with inequality, PDE and box constraints. There have been studies on the numerics of a Ginzburg–Landau regularization of the objective functional that can be shown to recover solutions with sharp transitions in the material distribution [19,20]. Notably, Garcke et al. [21] derived a posteriori error estimators designed to resolve the interfaces in the material distribution for the Navier–Stokes extension to the Borrvall–Petersson problem. As far as we are aware, there exist only a couple of results

* Corresponding author.

E-mail addresses: ioannis.papadopoulos@maths.ox.ac.uk (I.P.A. Papadopoulos), andre.suli@maths.ox.ac.uk (E. Süli).

¹ I. P. is supported by the EPSRC Centre for Doctoral Training in Partial Differential Equations: Analysis and Applications [Grant Number EP/L015811/1], the EPSRC grant Spectral element methods for fractional differential equations, with applications in applied analysis and medical imaging [Grant Number EP/T022132/1] and The MathWorks, Inc.

dealing with weak(-*) convergence of discretized solutions, as the mesh size tends to zero, to solutions of the Borrvall–Petersson problem on the whole domain [6,22]. Moreover, there have been no results concerning strong convergence nor the convergence to *all* the isolated minimizers of the problem.

In the original paper [6], it is shown that a minimizing velocity and material distribution to the optimization problem exist [6, Th. 3.1]; however, the minimizer is not necessarily unique [6, Sec. 4.5]. It is also shown that there exist finite element solutions that converge to a minimizer of the problem [6, Th. 3.2]. The proven convergence is weak in the approximation of the velocity and weak-* in the material distribution, with no results for the pressure. In addition, Borrvall and Petersson show that the approximation of material distribution strongly converges to a solution in $L^s(\Omega_b)$, $s \in [1, \infty)$, where Ω_b is any measurable subset of Ω in which the material distribution that solves the infinite-dimensional problem is equal to zero or one a.e. [6, Sec. 3.3]. Weak-* convergence permits large oscillations in the material distribution, called checkerboarding, which could occur in areas where the material distribution is not zero or one under the current results. However, in practice, checkerboarding is not observed in these regions. Since there can be multiple solutions, the nature of the convergence is ambiguous. In particular, it is not clear if there are sequences of finite element solutions converging to every solution of the infinite-dimensional problem.

Our goal is to extend and refine the analysis of Borrvall and Petersson. We show that, given an isolated minimizer to the infinite-dimensional problem, there exists a sequence of finite element solutions, satisfying the necessary first-order optimality conditions, that strongly converges. In particular, we strengthen the convergence from weak convergence in $H^1(\Omega)^d$ to strong convergence in $H^1(\Omega)^d$ for the velocity, and from weak-* convergence in $L^\infty(\Omega)$ to strong convergence in $L^s(\Omega)$, $s \in [1, \infty)$ for the material distribution. Moreover, in the case of a homogeneous Dirichlet boundary condition, we show that the material distribution is weakly differentiable inside any compact subset of the support of the velocity; more specifically $\rho \in H^1(U_\theta)$, for any $\theta > 0$, where U_θ is any measurable subset of Ω in which $|\mathbf{u}|^2 \geq \theta > 0$ a.e. in U_θ . This analysis confirms that checkerboarding cannot occur under mild assumptions on the model. Hence, isolated minimizers of the problem can be well approximated using the finite element method. We conclude with a numerical investigation into the convergence of the finite element solutions.

By first considering the optimization problem, we derive necessary first-order optimality conditions in Section 2. By construction, the generalized Stokes equations are satisfied but we also show that the material distribution satisfies a variational inequality. In Section 3, we show that under moderate assumptions, the material distribution is weakly differentiable in the case of a homogeneous Dirichlet boundary condition for the velocity. We tackle the issue of multiple local minima in Section 4, by considering closed balls around isolated local minimizers. In that section, we also prove that for each isolated minimizer there exists a sequence of finite element solutions to the discretized first-order optimality conditions, which strongly converges to the solution of the infinite-dimensional problem. In Section 5, we computationally investigate the convergence of sequences of finite element approximations to the respective solutions of the infinite-dimensional problem.

2. Existence and necessary first-order optimality conditions

The topology optimization problem of Borrvall and Petersson [6] is as follows: find the velocity, \mathbf{u} , and the material distribution, ρ , that solve the minimization problem

$$\min_{(\mathbf{v}, \eta) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma} J(\mathbf{v}, \eta), \quad (\text{BP})$$

where,

$$\begin{aligned} J(\mathbf{v}, \eta) &:= \frac{1}{2} \int_{\Omega} (\alpha(\eta)|\mathbf{v}|^2 + \nu|\nabla \mathbf{v}|^2 - 2\mathbf{f} \cdot \mathbf{v}) \, dx, \\ H_{\mathbf{g}}^1(\Omega)^d &:= \{\mathbf{v} \in H^1(\Omega)^d : \mathbf{v}|_{\partial\Omega} = \mathbf{g} \text{ on } \partial\Omega\}, \\ H_{\mathbf{g}, \text{div}}^1(\Omega)^d &:= \{\mathbf{v} \in H_{\mathbf{g}}^1(\Omega)^d : \text{div}(\mathbf{v}) = 0 \text{ a.e. in } \Omega\}, \\ C_\gamma &:= \left\{ \eta \in L^\infty(\Omega) : 0 \leq \eta \leq 1 \text{ a.e., } \int_{\Omega} \eta \, dx \leq \gamma|\Omega| \right\}, \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain with dimension $d \in \{2, 3\}$, $\mathbf{f} \in L^2(\Omega)^d$, ν is the (constant) viscosity, and $\gamma \in (0, 1)$ is the volume fraction. Here, α is the inverse permeability, modeling the influence of the material distribution on the flow. For values of ρ close to one, $\alpha(\rho)$ is small permitting fluid flow; for small values of ρ , $\alpha(\rho)$ is very large, restricting fluid flow. The function α is assumed to have the following properties:

- (A1) $\alpha : [0, 1] \rightarrow [\underline{\alpha}, \bar{\alpha}]$ with $0 \leq \underline{\alpha}$ and $\bar{\alpha} < \infty$;
- (A2) α is convex and monotonically decreasing;
- (A3) $\alpha(0) = \bar{\alpha}$ and $\alpha(1) = \underline{\alpha}$,

generating an operator also denoted $\alpha : C_\gamma \rightarrow L^\infty(\Omega; [\underline{\alpha}, \bar{\alpha}])$. Typically, in the literature α takes the form [6]

$$\alpha(\rho) = \bar{\alpha} \left(1 - \frac{\rho(q+1)}{\rho+q} \right), \quad (2.1)$$

where $q > 0$ is a penalty parameter, so that $\lim_{q \rightarrow \infty} \alpha(\rho) = \bar{\alpha}(1 - \rho)$. Furthermore, $|\partial\Omega$ is to be understood in the boundary trace sense [23, Ch. 5.5], $\mathbf{g} \in H^{1/2}(\partial\Omega)^d$, $\mathbf{g} = \mathbf{0}$ on $\Gamma \subset \partial\Omega$, with $\mathcal{H}^{d-1}(\Gamma) > 0$, i.e. Γ has nonzero $(d-1)$ -dimensional Hausdorff measure. Hence, the Poincaré inequality holds with constant c_p such that $\|\mathbf{v}\|_{L^2(\Omega)} \leq c_p \|\nabla \mathbf{v}\|_{L^2(\Omega)}$ for all $\mathbf{v} \in H^1(\Omega)^d$ with $\mathbf{v}|_\Gamma = \mathbf{0}$.

Remark 1. The integral in (BP) is well defined. Indeed, since α is assumed to be convex it is Borel measurable; also since $\rho \in C_\gamma$ is Lebesgue measurable, the composition $\alpha(\rho) : \Omega \rightarrow [\underline{\alpha}, \bar{\alpha}]$ is Lebesgue measurable.

The next theorem is due to Borrvall and Petersson [6, Th. 3.1].

Theorem 1. Suppose that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain, $d \in \{2, 3\}$, and α satisfies properties (A1)–(A3). Suppose in addition that $\alpha \in C^1([0, 1]; [\underline{\alpha}, \bar{\alpha}])$. Then, there exists a pair $(\mathbf{u}, \rho) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma$ that locally minimizes J , as defined in (BP).

Remark 2. Although a solution exists, it is not necessarily unique, as observed in the numerical examples in Section 5, since the optimization problem is nonconvex. The nonconvexity is caused by the term $\alpha(\rho)|\mathbf{u}|^2$ in (BP). A rough argument examines the second partial Fréchet derivative of J with respect to \mathbf{u} and ρ (assuming that it exists), i.e., for suitable variations \mathbf{v}, η ,

$$\langle J''_{\mathbf{u}, \rho}(\mathbf{u}, \rho), \mathbf{v}, \eta \rangle = \int_{\Omega} \alpha'(\rho) \eta (\mathbf{u} \cdot \mathbf{v}) \, dx. \quad (2.2)$$

Note that (A2) implies that $\alpha'(\rho) < 0$ a.e. and the box constraints enforce $\rho \geq 0$ a.e. Thus, given a pair (\mathbf{u}, ρ) , such that $\mathbf{u} \geq \mathbf{0}$ a.e., \mathbf{u}, ρ are nonzero functions, and nonzero variations (\mathbf{v}, η) , such that $\mathbf{v} \geq \mathbf{0}$ a.e. and $\eta > 0$ a.e., we see that $\langle J''_{\mathbf{u}, \rho}(\mathbf{u}, \rho), \mathbf{v}, \eta \rangle < 0$. Moreover, if $2\langle J''_{\mathbf{u}, \rho}(\mathbf{u}, \rho), \mathbf{v}, \eta \rangle + \langle J''_{\mathbf{u}, \mathbf{u}}(\mathbf{u}, \rho), \mathbf{v}, \mathbf{v} \rangle + \langle J''_{\rho, \rho}(\mathbf{u}, \rho), \eta, \eta \rangle < 0$, then the optimization problem is not convex.

Proposition 1. Suppose α also satisfies

(A4) α is twice continuously differentiable.

Then $J : H^1(\Omega)^d \times L^s(\Omega) \rightarrow \mathbb{R}$ is partially Fréchet differentiable with respect to \mathbf{u} and partially Fréchet differentiable with respect to ρ , where $1 < s \leq \infty$ in two dimensions and $3/2 \leq s \leq \infty$ in three dimensions. Moreover, for all $\mathbf{v} \in H_0^1(\Omega)^d$ and $\eta \in C_\gamma$ we have that

$$\langle J'_\mathbf{u}(\mathbf{u}, \rho), \mathbf{v} \rangle = \int_{\Omega} \alpha(\rho) \mathbf{u} \cdot \mathbf{v} + \nu \nabla \mathbf{u} : \nabla \mathbf{v} - \mathbf{f} \cdot \mathbf{v} \, dx, \quad (2.3)$$

$$\langle J'_\rho(\mathbf{u}, \rho), \eta - \rho \rangle = \frac{1}{2} \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\eta - \rho) \, dx, \quad (2.4)$$

where $J'_\mathbf{u}(\mathbf{u}, \rho)$ denotes the Fréchet derivative of J with respect to \mathbf{u} and $J'_\rho(\mathbf{u}, \rho)$ denotes the Fréchet derivative of J with respect to ρ .

Proof. Proposition 1 follows from the definition of Fréchet differentiability, the mean value inequality, the dominated convergence theorem, and the Sobolev embedding theorem [24, Prop. 2.3]. \square

Remark 3. It can be checked that if α is $(n+1)$ -times continuously differentiable then J is n -times Fréchet differentiable with respect to \mathbf{u} and ρ .

The following proposition is the main result of this section. We show that if (\mathbf{u}, ρ) is a minimizer of the optimization problem (BP), then the minimizer also satisfies first-order optimality conditions consisting of two equations and a variational inequality.

Proposition 2. Suppose that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain, with $d \in \{2, 3\}$, and α satisfies properties (A1)–(A4). Consider a local or global minimizer $(\mathbf{u}, \rho) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma$ of (BP). Then, there exists a unique Lagrange multiplier $p \in L_0^2(\Omega)$ such that the following necessary first-order optimality conditions hold:

$$a_\rho(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = l_f(\mathbf{v}) \quad \text{for all } \mathbf{v} \in H_0^1(\Omega)^d, \quad (\text{FOC1})$$

$$b(\mathbf{u}, q) = 0 \quad \text{for all } q \in L_0^2(\Omega), \quad (\text{FOC2})$$

$$c_\mathbf{u}(\rho, \eta - \rho) \geq 0 \quad \text{for all } \eta \in C_\gamma, \quad (\text{FOC3})$$

where

$$L_0^2(\Omega) := \left\{ q \in L^2(\Omega) : \int_{\Omega} q \, dx = 0 \right\},$$

and

$$\begin{aligned} a_\rho(\mathbf{u}, \mathbf{v}) &:= \int_{\Omega} [\alpha(\rho)\mathbf{u} \cdot \mathbf{v} + \nu \nabla \mathbf{u} : \nabla \mathbf{v}] \, dx, \quad l_f(\mathbf{v}) := \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, dx, \\ b(\mathbf{v}, q) &:= - \int_{\Omega} q \operatorname{div}(\mathbf{v}) \, dx, \quad c_u(\rho, \eta) := \frac{1}{2} \int_{\Omega} \alpha'(\rho) \eta |\mathbf{u}|^2 \, dx. \end{aligned}$$

Proof. We will first show that (FOC1)–(FOC2) are satisfied by generalizing arguments, used for the Stokes system with a homogeneous Dirichlet boundary condition, found in [25]. For ease of notation we define $X_g := H_g^1(\Omega)^d$, $X_0 := H_0^1(\Omega)^d$, $V_g := H_{g,\operatorname{div}}^1(\Omega)^d$, $V_0 := H_{0,\operatorname{div}}^1(\Omega)^d$ and $M := L_0^2(\Omega)$. The respective dual spaces of X_0 , V_0 and M are denoted with $*$. We also define the associated operators, $A \in \mathcal{L}(X_g, X_0^*)$, $B \in \mathcal{L}(X_g, M)$ and $B_0 \in \mathcal{L}(X_0, M)$ by

$$\langle \mathbf{A}\mathbf{u}, \mathbf{v} \rangle := a_\rho(\mathbf{u}, \mathbf{v}), \quad \langle B\mathbf{w}, q \rangle := b(\mathbf{w}, q), \quad \text{and} \quad \langle B_0\mathbf{v}, q \rangle := b(\mathbf{v}, q). \quad (2.5)$$

We note that $\ker(B_0) = V_0$. From Theorem 1, we know that there exists a pair $(\mathbf{u}, \rho) \in V_g \times C_\gamma$ that is a local minimizer for (BP). For any given $\mathbf{v} \in V_0$, we see that $\mathbf{u} + t\mathbf{v} \in V_g$, $t \in \mathbb{R}$. If $(\mathbf{u}, \rho) \in V_g \times C_\gamma$ is a minimizer, then, by definition, there exists an $r > 0$ such that, for any $(\mathbf{w}, \eta) \in V_g \times C_\gamma$, $(\mathbf{w}, \eta) \neq (\mathbf{u}, \rho)$ that satisfies

$$\|\mathbf{u} - \mathbf{w}\|_{H^1(\Omega)} + \|\rho - \eta\|_{L^\infty(\Omega)} < r \quad (2.6)$$

we have that $J(\mathbf{u}, \rho) \leq J(\mathbf{w}, \eta)$. Hence, for any given $\mathbf{v} \in V_0$, if $0 < t < r/\|\mathbf{v}\|_{H^1(\Omega)}$, the following inequality holds

$$\frac{1}{t}(J(\mathbf{u} + t\mathbf{v}, \rho) - J(\mathbf{u}, \rho)) \geq 0. \quad (2.7)$$

By Proposition 1, J is Fréchet differentiable, and therefore also Gateaux differentiable, with respect to \mathbf{u} . Hence as $t \rightarrow 0_+$, we see that

$$\langle J'_u(\mathbf{u}, \rho), \mathbf{v} \rangle \geq 0 \quad \text{for all } \mathbf{v} \in V_0. \quad (2.8)$$

By considering the same reasoning with $t < 0$, we deduce that

$$\langle J'_u(\mathbf{u}, \rho), \mathbf{v} \rangle = 0 \quad \text{for all } \mathbf{v} \in V_0. \quad (2.9)$$

From Proposition 1, we know that $J'_u(\mathbf{u}, \rho) = \mathbf{A}\mathbf{u} - \mathbf{f}$ and hence $\mathbf{A}\mathbf{u} - \mathbf{f} \in V_0^\circ$ where

$$V_0^\circ := (\ker(B_0))^\circ = \{h \in X_0^* : \langle h, \mathbf{v} \rangle = 0 \text{ for all } \mathbf{v} \in V_0\}. \quad (2.10)$$

We know that the operator B_0 satisfies the following equivalent version of the inf-sup condition [26, Ch. 1, Sec. 4.1, Lem. 4.1]:

$$\text{there exists a } \beta > 0 \text{ such that, for all } q \in M, \quad \|B_0^*q\|_{X_0^*} \geq \beta \|q\|_M, \quad (2.11)$$

where B_0^* is the dual operator of B_0 , defined by $\langle \mathbf{v}, B_0^*q \rangle = \langle B_0\mathbf{v}, q \rangle$. This implies that B_0^* is injective (and therefore bijective) from M into $\operatorname{Im} B_0^*$. Furthermore, it also implies that $(B_0^*)^{-1}$ is continuous. Consider $\mathbf{f} \in \operatorname{Im} B_0^*$; then, there exists a $q \in M$ such that $\mathbf{f} = B_0^*q$ and

$$\|(B_0^*)^{-1}\mathbf{f}\|_M \leq \frac{1}{\beta} \|\mathbf{f}\|_{X_0^*}. \quad (2.12)$$

Therefore, $\operatorname{Im} B_0^*$ is closed.

Since $\operatorname{Im} B_0^*$ is closed, by Banach's closed range theorem, we know that $\operatorname{Im} B_0^* = (\ker(B_0))^\circ = V_0^\circ$. Hence, since $\mathbf{A}\mathbf{u} - \mathbf{f} \in V_0^\circ$, there exists a $p \in M$ such that

$$\mathbf{A}\mathbf{u} + B_0^*p = \mathbf{f}. \quad (2.13)$$

Since B_0^* is injective, p is also unique. Since $\mathbf{u} \in V_g$, we have that $B\mathbf{u} = 0$. Hence (FOC1) and (FOC2) hold.

We will now show that (FOC3) holds. We note that C_γ is a convex subset of a linear space. For any given $\zeta, \eta \in C_\gamma$ and $t \in [0, 1]$, we therefore have that $\zeta + t(\eta - \zeta) \in C_\gamma$. Since (\mathbf{u}, ρ) is a local minimizer, it follows that for each $\eta \in C_\gamma$, if $0 < t < r/\|\eta - \rho\|_{L^\infty(\Omega)}$, with r as in (2.6), then

$$\frac{1}{t}(J(\mathbf{u}, \rho + t(\eta - \rho)) - J(\mathbf{u}, \rho)) \geq 0. \quad (2.14)$$

From Proposition 1, we know that J is Fréchet differentiable, and therefore also Gateaux differentiable, with respect to ρ . Hence, by taking the limit as $t \rightarrow 0$, we see that

$$c_u(\rho, \eta - \rho) = \langle J'_\rho(\mathbf{u}, \rho), \eta - \rho \rangle \geq 0 \quad \text{for all } \eta \in C_\gamma. \quad (2.15)$$

Therefore (FOC3) holds. \square

The following lemma will be used in the proof of the next proposition.

Lemma 2. Consider a nonzero function $\eta \in C_\gamma$ and the measurable non-empty set $E \subset \subset \text{supp}(\eta)$, where $\subset \subset$ denotes that the containment is compact and supp denotes the support of a function, i.e. $\eta > 0$ a.e. in E . Then, there exists an $\epsilon' > 0$ such that, for all $\epsilon \in (0, \epsilon']$, there exists a set $E_\epsilon \subseteq E$, $|E_\epsilon| > 0$ where $\eta > \epsilon$ a.e. in E_ϵ .

Proof. For a contradiction, suppose that there exists no such ϵ' such that $E_{\epsilon'}$ exists. This implies that

$$\text{for all } n \geq 0, \quad |E \setminus \hat{E}_n| = 0, \quad (2.16)$$

where $\hat{E}_n := \{0 \leq \eta \leq 1/n \text{ a.e. in } E\}$. We see that $\emptyset = E \setminus \hat{E}_1 \subseteq E \setminus \hat{E}_2 \subseteq \dots \subseteq E \setminus \hat{E}_n \subseteq \dots$, i.e. $E \setminus \hat{E}_n$ is nondecreasing. Note that the limit of a nondecreasing sequence of sets (A_n) can be defined as $\lim_{n \rightarrow \infty} A_n := \cup_{n \geq 1} A_n$. By (2.16) we note that

$$\lim_{n \rightarrow \infty} |E \setminus \hat{E}_n| = 0. \quad (2.17)$$

Moreover,

$$\begin{aligned} \cup_{n=1}^{\infty} E \setminus \hat{E}_n &= \lim_{n \rightarrow \infty} E \setminus \{0 \leq \eta \leq 1/n \text{ a.e. in } E\} \\ &= E \setminus \{\eta = 0 \text{ a.e. in } E\} = E \setminus \emptyset = E. \end{aligned} \quad (2.18)$$

Now we see that

$$0 < |E| = |\cup_{n=1}^{\infty} E \setminus \hat{E}_n| = |\lim_{n \rightarrow \infty} E \setminus \hat{E}_n| = \lim_{n \rightarrow \infty} |E \setminus \hat{E}_n| = 0, \quad (2.19)$$

where the first equality follows from (2.18), the second equality follows from the definition of the limit of a nondecreasing sequence of sets, the third equality follows from the continuity of the Lebesgue measure, and the fourth equality follows from (2.17). (2.19) is a contradiction and, therefore, such an $\epsilon' > 0$ must exist. By choosing $E_\epsilon = E_{\epsilon'}$ for all $0 < \epsilon \leq \epsilon'$, we conclude that the statement holds for all $\epsilon \in (0, \epsilon']$. \square

In the result that follows, we are required to distinguish between different types of global and local minimizers.

Definition 1 (Strict Minimizer). Let Z be a Banach space and suppose that $z_0 \in Z$ is a local or global minimizer of the functional $J : Z \rightarrow \mathbb{R}$. We say that z_0 is a strict minimizer if there exists an open neighborhood $E \subset Z$ of z_0 such that $J(z_0) < J(z)$ for all $z \neq z_0, z \in E$.

Definition 2 (Isolated Minimizer). Let Z be a Banach space and suppose that $z_0 \in Z$ is a local or global minimizer of the functional $J : Z \rightarrow \mathbb{R}$. We say that z_0 is isolated if there exists an open neighborhood $E \subset Z$ of z_0 such that there are no other minimizers contained in E .

Remark 4. If z is an isolated minimizer, then it is also a strict minimizer.

The following proposition is a property of strict minimizers that will be useful for the numerical analysis of the finite element method.

Proposition 3. Suppose that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain, $d \in \{2, 3\}$, and α satisfies properties (A1)–(A4). Further assume that the minimizer $(\mathbf{u}, \rho) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma$ of (BP) is a strict minimizer. Then, $\text{supp}(\rho) \subseteq U$, where $U := \text{supp}(\mathbf{u})$.

Proof. By definition of a strict minimizer, there exists an $r > 0$ such that, for all $(\mathbf{w}, \eta) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma$, $(\mathbf{w}, \eta) \neq (\mathbf{u}, \rho)$ that satisfies

$$\|\mathbf{u} - \mathbf{w}\|_{H^1(\Omega)} + \|\rho - \eta\|_{L^\infty(\Omega)} < r,$$

we have that $J(\mathbf{u}, \rho) < J(\mathbf{w}, \eta)$. For a contradiction, suppose that there exists a set $E \subset \Omega$, $E \cap U = \emptyset$, of positive measure, where $\rho > 0$ a.e. in E . By Lemma 2, there exists an $\epsilon \in (0, r)$ such that there exists a set $E_\epsilon \subseteq E$, $|E_\epsilon| > 0$ where $\rho > \epsilon$ a.e. in E_ϵ . Define $\tilde{\rho}$ as

$$\tilde{\rho} := \begin{cases} \rho & \text{a.e. in } \Omega \setminus E_\epsilon, \\ \rho - \epsilon & \text{a.e. in } E_\epsilon. \end{cases} \quad (2.20)$$

As $\rho \in C_\gamma$, also $\tilde{\rho} \in C_\gamma$. We note that $\|\rho - \tilde{\rho}\|_{L^\infty(\Omega)} = \|\epsilon\|_{L^\infty(E_\epsilon)} < r$ and, therefore, $(\mathbf{u}, \tilde{\rho})$ lies inside the minimizing neighborhood of the (\mathbf{u}, ρ) . However, $J(\mathbf{u}, \tilde{\rho}) = J(\mathbf{u}, \rho)$ as ρ and $\tilde{\rho}$ only differ on the set E_ϵ , but $\mathbf{u} = \mathbf{0}$ a.e. in $E_\epsilon \subseteq E$ by assumption. This contradicts the assertion that (\mathbf{u}, ρ) is a strict minimizer. \square

3. Regularity of ρ

In this section we show that $\rho \in C_\gamma$ possesses higher regularity in the case of a homogeneous Dirichlet boundary condition on \mathbf{u} and if α satisfies a stronger (but not restrictive) convexity assumption.

Theorem 3 (Regularity of ρ). *Suppose that the domain $\Omega \subset \mathbb{R}^d$ is bounded, the boundary is Lipschitz, and that the data $\mathbf{g} = \mathbf{0}$ on $\partial\Omega$. Consider a local or global minimizer, $(\mathbf{u}, \rho) \in H_{0,\text{div}}^1(\Omega)^d \times C_\gamma$, of (BP) such that \mathbf{u} is not the zero function and there exists a closed subset $\bar{U}_\theta \subset \Omega$ with non-empty interior on which $|\mathbf{u}|^2$ is bounded below by a positive constant, $\theta > 0$. Suppose that (A1)–(A4) hold and that $\alpha \in C^2([0, 1])$ is strongly convex, i.e.,*

(A5) *There exists a constant $\alpha''_{\min} > 0$ such that $\alpha''(y) \geq \alpha''_{\min} > 0$ for all $y \in [0, 1]$.*

Consider the (non-empty) open interior $U_\theta \subset \bar{U}_\theta$. Then, $\nabla \rho$ exists in U_θ and $\rho \in C_\gamma \cap H^1(U_\theta)$.

Remark 5. The assumption (A5) excludes the case where α is linear. This is consistent with previous theory, as Borrvall and Petersson [6] showed that if α is linear, then there exists a minimizer (\mathbf{u}, ρ) where ρ is a 0–1 solution (a linear combination of Heaviside functions) and thus $\rho \notin H^1(\Omega)$ due to the jumps. However, the assumptions (A1)–(A5) do include (2.1), where the lower bound in (A5) is $\alpha''_{\min} = 2\bar{\alpha}q/(q+1)^2$. We see that this lower bound degrades to zero as $q \rightarrow \infty$. As previously noted, the limit $q \rightarrow \infty$ coincides with $\alpha(\rho) \rightarrow \bar{\alpha}(1 - \rho)$, which is a linear function.

Proof of Theorem 3. Let ∂_{x_k} denote the partial derivative with respect to x_k . If we can bound the L^2 -norm of the difference quotients of ρ , in all coordinate directions in U_θ , above by constants independent of h , then, by taking the weak limit, we deduce that $\partial_{x_k}\rho$ exists as an element of $L^2(U_\theta)$ for $1 \leq k \leq d$.

The variational inequality on ρ states that

$$\frac{1}{2} \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\eta - \rho) \, dx \geq 0. \quad (3.1)$$

We define $U \subset \Omega$ as $U := \text{supp}(\mathbf{u})$ and fix an open, bounded and connected domain $\hat{\Omega}$ such that $\hat{\Omega} = \Omega$ if $U \subset\subset \Omega$ and $\Omega \subset\subset \hat{\Omega}$ otherwise. In the case where U is not a compact subset of Ω , we extend \mathbf{u} and ρ by zero to the whole of \mathbb{R}^d . Since the trace of \mathbf{u} is zero on the boundary, the extension of \mathbf{u} by zero lives in $H^1(\hat{\Omega})^d$. Let $0 < |h| < (1/2)\text{dist}(U, \partial\hat{\Omega})$ and choose $k \in \{1, \dots, d\}$. We define ρ^h as

$$\rho^h(x) = \begin{cases} \rho(x + he_k) & \text{for } x \in \Omega - he_k, \\ 0 & \text{for } x \in \mathbb{R}^d \setminus (\Omega - he_k). \end{cases}$$

Analogously, we define \mathbf{u}^h as

$$\mathbf{u}^h(x) = \begin{cases} \mathbf{u}(x + he_k) & \text{for } x \in \Omega - he_k, \\ \mathbf{0} & \text{for } x \in \mathbb{R}^d \setminus (\Omega - he_k). \end{cases}$$

We define the difference quotient, D_k^h , in the k th coordinate direction, as

$$D_k^h \rho(x) = \frac{\rho(x + he_k) - \rho(x)}{h}, \quad h \in \mathbb{R} \setminus \{0\}, \quad x \in \hat{\Omega}.$$

Let $\eta = (\rho^h + \rho^{-h})/2$. We note that $\eta \in C_\gamma$, since

$$0 \leq \frac{1}{2} \rho^h \leq \frac{1}{2} \text{ a.e. in } \Omega, \quad \text{and} \quad 0 \leq \frac{1}{2} \rho^{-h} \leq \frac{1}{2} \text{ a.e. in } \Omega,$$

which implies that $0 \leq \eta \leq 1$ a.e. in Ω and

$$\begin{aligned} \int_{\Omega} \eta \, dx &= \frac{1}{2} \int_{\Omega} \rho^h + \rho^{-h} \, dx \\ &= \frac{1}{2} \int_{\Omega - he_k \cap \Omega} \rho \, dx + \frac{1}{2} \int_{\Omega + he_k \cap \Omega} \rho \, dx \leq \int_{\Omega} \rho \, dx \leq \gamma |\Omega|. \end{aligned}$$

If we multiply (3.1) through by 4 and divide by h^2 we see that

$$\frac{1}{h^2} \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\rho^h + \rho^{-h} - 2\rho) \, dx \geq 0. \quad (3.2)$$

We note that,

$$D_k^{-h}(D_k^h \rho) = \frac{\frac{\rho - \rho^{-h}}{h} - \frac{\rho^h - \rho}{h}}{-h} = \frac{\rho^h + \rho^{-h} - 2\rho}{h^2}.$$

Hence, because \mathbf{u} is zero outside of Ω , (3.2) is equivalent to

$$\int_{\hat{\Omega}} \alpha'(\rho) |\mathbf{u}|^2 (D_k^{-h}(D_k^h \rho)) \, dx \geq 0. \quad (3.3)$$

In order to obtain a first-order difference quotient, we will perform the finite difference analogue of integration by parts to shift the D_k^{-h} operator from $D_k^h \rho$ to $\alpha'(\rho) |\mathbf{u}|^2$. We note that, by definition, the left-hand side of (3.3) is equal to

$$-\frac{1}{h} \int_{\hat{\Omega}} (\alpha'(\rho) |\mathbf{u}|^2)(x) ((D_k^h \rho)(x - h e_k) - (D_k^h \rho)(x)) \, dx, \quad (3.4)$$

which by a change of variables is equal to

$$-\frac{1}{h} \left(\int_{\hat{\Omega} - h e_k} (\alpha'(\rho) |\mathbf{u}|^2)(x + h e_k) (D_k^h \rho)(x) \, dx - \int_{\hat{\Omega}} (\alpha'(\rho) |\mathbf{u}|^2)(x) (D_k^h \rho)(x) \, dx \right).$$

We note that $U \subset \subset \hat{\Omega}$ and $|h| < (1/2) \text{dist}(U, \partial \hat{\Omega})$, which implies that $U \subset \subset \hat{\Omega} - h e_k$. Therefore,

$$\begin{aligned} & \int_{\hat{\Omega} - h e_k} (\alpha'(\rho) |\mathbf{u}|^2)(x + h e_k) (D_k^h \rho)(x) \, dx \\ &= \int_{U - h e_k} (\alpha'(\rho) |\mathbf{u}|^2)(x + h e_k) (D_k^h \rho)(x) \, dx \\ &= \int_{\hat{\Omega}} (\alpha'(\rho) |\mathbf{u}|^2)(x + h e_k) (D_k^h \rho)(x) \, dx. \end{aligned} \quad (3.5)$$

Therefore, from (3.3)–(3.5) we see that

$$\int_{\hat{\Omega}} D_k^h (\alpha'(\rho) |\mathbf{u}|^2) (D_k^h \rho) \, dx \leq 0. \quad (3.6)$$

Now we wish to rewrite $D_k^h (\alpha'(\rho) |\mathbf{u}|^2)$ in a form that we can decouple from $D_k^h \rho$ in order to be able to bound (3.6) above and below. Now,

$$\begin{aligned} D_k^h (\alpha'(\rho) |\mathbf{u}|^2)(x) &= \frac{1}{h} (\alpha'(\rho(x + h e_k)) |\mathbf{u}(x + h e_k)|^2 - \alpha'(\rho(x)) |\mathbf{u}(x)|^2) \\ &= \frac{1}{2h} (\alpha'(\rho(x + h e_k)) (|\mathbf{u}(x + h e_k)|^2 - |\mathbf{u}(x)|^2)) \\ &\quad + \frac{1}{2h} (\alpha'(\rho(x)) (|\mathbf{u}(x + h e_k)|^2 - |\mathbf{u}(x)|^2)) \\ &\quad + \frac{1}{2h} (|\mathbf{u}(x + h e_k)|^2 (\alpha'(\rho(x + h e_k)) - \alpha'(\rho(x)))) \\ &\quad + \frac{1}{2h} (|\mathbf{u}(x)|^2 (\alpha'(\rho(x + h e_k)) - \alpha'(\rho(x)))) \\ &= \frac{1}{2} (\alpha'(\rho^h) + \alpha'(\rho)) D_k^h (|\mathbf{u}|^2) + \frac{1}{2} (|\mathbf{u}^h|^2 + |\mathbf{u}|^2) D_k^h (\alpha'(\rho)). \end{aligned}$$

Therefore, from (3.6) we see that

$$\begin{aligned} & \int_{\hat{\Omega}} \left[\frac{1}{2} (|\mathbf{u}^h|^2 + |\mathbf{u}|^2) D_k^h (\alpha'(\rho)) \right. \\ & \quad \left. + \frac{1}{2} (\alpha'(\rho^h) + \alpha'(\rho)) D_k^h |\mathbf{u}|^2 \right] D_k^h (\rho) \, dx \leq 0. \end{aligned} \quad (3.7)$$

Now,

$$\begin{aligned} & \frac{1}{2} (|\mathbf{u}^h|^2 + |\mathbf{u}|^2) D_k^h (\alpha'(\rho)) + \frac{1}{2} (\alpha'(\rho^h) + \alpha'(\rho)) D_k^h |\mathbf{u}|^2 \\ &= \frac{1}{h} \int_0^1 \frac{d}{ds} \left[\alpha' (s \rho^h + (1-s) \rho) \frac{1}{2} (|\mathbf{u}^h|^2 + |\mathbf{u}|^2) \right. \\ & \quad \left. + \frac{1}{2} (\alpha'(\rho^h) + \alpha'(\rho)) |s \mathbf{u}^h + (1-s) \mathbf{u}|^2 \right] ds \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{h} \underbrace{\int_0^1 [\alpha''(s\rho^h + (1-s)\rho)] ds}_{=:A} \frac{1}{2} (|\mathbf{u}^h|^2 + |\mathbf{u}|^2) (\rho^h - \rho) \\
&\quad + \frac{1}{2h} (\alpha'(\rho^h) + \alpha'(\rho)) \underbrace{\int_0^1 [2(\mathbf{su}^h + (1-s)\mathbf{u})] ds}_{=:B} \cdot (\mathbf{u}^h - \mathbf{u}).
\end{aligned}$$

Hence from (3.7) we find that

$$\frac{1}{2} \int_{\hat{\Omega}} A(|\mathbf{u}^h|^2 + |\mathbf{u}|^2) |D_k^h \rho|^2 + (\alpha'(\rho^h) + \alpha'(\rho)) \mathbf{B} \cdot (D_k^h \mathbf{u}) D_k^h \rho \, dx \leq 0. \quad (3.8)$$

Subtracting the second term on the left-hand side in (3.8) from both sides, taking absolute values on the right-hand side, using the Cauchy–Schwarz inequality and multiplying by 2, we see that

$$\int_{\hat{\Omega}} A(|\mathbf{u}^h|^2 + |\mathbf{u}|^2) |D_k^h \rho|^2 \, dx \leq \int_{\hat{\Omega}} |\mathbf{B}| |\alpha'(\rho^h) + \alpha'(\rho)| |D_k^h \mathbf{u}| |D_k^h \rho| \, dx. \quad (3.9)$$

Furthermore we note that $A \geq \alpha''_{\min}$ and

$$\mathbf{B} = \int_0^1 [2(\mathbf{su}^h + (1-s)\mathbf{u})] ds = 2 \left[\frac{s^2}{2} \mathbf{u}^h + \left(s - \frac{s^2}{2} \right) \mathbf{u} \right]_0^1 = \mathbf{u}^h + \mathbf{u}.$$

Hence, using Cauchy's inequality and Young's inequality, we see that

$$\begin{aligned}
&\alpha''_{\min} \int_{U-\text{he}_k} |\mathbf{u}^h|^2 |D_k^h \rho|^2 \, dx + \alpha''_{\min} \int_U |\mathbf{u}|^2 |D_k^h \rho|^2 \, dx \\
&\leq \int_{\hat{\Omega}} |\mathbf{u} + \mathbf{u}^h| |\alpha'(\rho^h) + \alpha'(\rho)| |D_k^h \mathbf{u}| |D_k^h \rho| \, dx \\
&\leq \int_{U-\text{he}_k} |\mathbf{u}^h| |\alpha'(\rho^h) + \alpha'(\rho)| |D_k^h \mathbf{u}| |D_k^h \rho| \, dx \\
&\quad + \int_U |\mathbf{u}| |\alpha'(\rho^h) + \alpha'(\rho)| |D_k^h \mathbf{u}| |D_k^h \rho| \, dx \\
&\leq \frac{\epsilon}{2} \int_{U-\text{he}_k} |\mathbf{u}^h|^2 |D_k^h \rho|^2 \, dx + \frac{\epsilon}{2} \int_U |\mathbf{u}|^2 |D_k^h \rho|^2 \, dx \\
&\quad + \frac{1}{2\epsilon} \int_{U-\text{he}_k} |\alpha'(\rho^h) + \alpha'(\rho)|^2 |D_k^h \mathbf{u}|^2 \, dx \\
&\quad + \frac{1}{2\epsilon} \int_U |\alpha'(\rho^h) + \alpha'(\rho)|^2 |D_k^h \mathbf{u}|^2 \, dx.
\end{aligned} \quad (3.10)$$

By fixing $\epsilon = \alpha''_{\min}$, from (3.10) we see that,

$$\begin{aligned}
&\frac{\alpha''_{\min}}{2} \int_U |\mathbf{u}|^2 |D_k^h \rho|^2 \, dx \\
&\leq \frac{\alpha''_{\min}}{2} \int_U |\mathbf{u}|^2 |D_k^h \rho|^2 \, dx + \frac{\alpha''_{\min}}{2} \int_{U-\text{he}_k} |\mathbf{u}^h|^2 |D_k^h \rho|^2 \, dx \\
&\leq \frac{1}{\alpha''_{\min}} \int_{\hat{\Omega}} |\alpha'(\rho^h) + \alpha'(\rho)|^2 |D_k^h \mathbf{u}|^2 \, dx.
\end{aligned} \quad (3.11)$$

Now $|\alpha'(\rho^h) + \alpha'(\rho)|^2$ is bounded above by $4 \sup_{\zeta \in C_\gamma} |\alpha'(\zeta)|^2$ which is independent of h . Consider a set $\tilde{\Omega} \subset \mathbb{R}^d$ such that $\hat{\Omega} \subset \subset \tilde{\Omega}$. We note that $\mathbf{u} \in H^1(\tilde{\Omega})^d$. By applying Theorem 3 in [23, pg. 294], we see that

$$\begin{aligned}
\int_U |\mathbf{u}|^2 |D_k^h \rho|^2 \, dx &\leq \frac{\tilde{C}(\Omega) \sup_{\zeta \in C_\gamma} |\alpha'(\zeta)|^2}{(\alpha''_{\min})^2} \|\nabla \mathbf{u}\|_{L^2(\tilde{\Omega})}^2 \\
&\leq \frac{\hat{C}(\Omega) \sup_{\zeta \in C_\gamma} |\alpha'(\zeta)|^2}{(\alpha''_{\min})^2} \|\nabla \mathbf{u}\|_{L^2(\Omega)}^2 \leq C < \infty,
\end{aligned} \quad (3.12)$$

where \tilde{C} , \hat{C} and C are constants. The bound is independent of h and k . Because, by hypothesis, there exists a subset $U_\theta \subset \Omega$ such that, $|\mathbf{u}|^2 \geq \theta > 0$ a.e. in U_θ , we see from (3.12) that, because $U_\theta \subset U = \text{supp}(\mathbf{u})$, also

$$\theta \int_{U_\theta} |D_k^h \rho|^2 dx \leq \int_{U_\theta} |\mathbf{u}|^2 |D_k^h \rho|^2 dx \leq \int_U |\mathbf{u}|^2 |D_k^h \rho|^2 dx \leq C. \quad (3.13)$$

Estimate (3.13) implies that

$$\sup_h \|D_k^h \rho\|_{L^2(U_\theta)} < \infty. \quad (3.14)$$

From (3.14) we see that there exists a function $\eta_k \in L^2(U_\theta)$ and a subsequence $h_i \rightarrow 0$ such that,

$$D_k^{h_i} \rho \rightharpoonup \eta_k \text{ weakly in } L^2(U_\theta).$$

Finally, we wish to identify η_k with $\partial_{x_k} \rho$. First choose any smooth and compactly supported function, $\phi \in C_c^\infty(U_\theta)$. We note that

$$\int_{U_\theta} \rho \partial_{x_k} \phi dx \leq C \|\rho\|_{L^\infty(U_\theta)} \|\phi\|_{W^{1,\infty}(U_\theta)} < \infty.$$

Since $\partial_{x_k} \phi$ is compactly supported in U_θ , it follows that

$$\int_{U_\theta} \rho \partial_{x_k} \phi dx = \int_{\hat{\Omega}} \rho \partial_{x_k} \phi dx.$$

Hence

$$\begin{aligned} \int_{U_\theta} \rho \partial_{x_k} \phi dx &= \lim_{h_i \rightarrow 0} \int_{\hat{\Omega}} \rho D_k^{-h_i} \phi dx \\ &= - \lim_{h_i \rightarrow 0} \int_{\hat{\Omega}} (D_k^{h_i} \rho) \phi dx = - \lim_{h_i \rightarrow 0} \int_{U_\theta} (D_k^{h_i} \rho) \phi dx = - \int_{U_\theta} \eta_k \phi dx. \end{aligned}$$

Hence $\eta_k = \partial_{x_k} \rho$ a.e. in U_θ for $k = 1, \dots, d$. Therefore, from (3.13) we see by weak lower semicontinuity that

$$\int_{U_\theta} |\partial_{x_k} \rho|^2 dx \leq C(\Omega, \sup_{\zeta \in C_\gamma} |\alpha'(\zeta)|^2, \alpha''_{\min}, \theta), \quad (3.15)$$

for some constant C . We conclude that $\rho \in H^1(U_\theta) \cap C_\gamma$, $\theta > 0$. \square

Remark 6. The assumption that the boundary datum $\mathbf{g} = \mathbf{0}$ on $\partial\Omega$ is required so as to expand the domain of integration from Ω in (3.2) to $\hat{\Omega}$ in (3.3) in order to perform the finite difference analogue of integration by parts in (3.4). A homogeneous Dirichlet boundary condition on \mathbf{u} is rarely imposed in practice. However, we observe during numerical experiments that ρ possesses additional regularity in the case of inhomogeneous Dirichlet boundary conditions, and we hypothesize that the results can be generalized to that case.

4. Finite element approximation

We will be approximately solving (FOC1)–(FOC3) by approximating the solutions of the infinite-dimensional problem with finite element functions. Borrvall and Petersson [6, Sec. 3.3] considered a piecewise constant finite element approximation of the material distribution coupled with an inf–sup stable quadrilateral finite element approximation of the velocity and the pressure. They showed that such approximations of the velocity and material distribution converge to an unspecified solution (\mathbf{u}, ρ) of (BP) in the following sense [6]:

$$\begin{aligned} \mathbf{u}_h &\rightharpoonup \mathbf{u} \text{ weakly in } H^1(\Omega)^d, \\ \rho_h &\overset{*}{\rightharpoonup} \rho \text{ weakly-}^* \text{ in } L^\infty(\Omega), \\ \rho_h &\rightarrow \rho \text{ strongly in } L^s(\Omega_b), \quad s \in [1, \infty), \end{aligned}$$

where Ω_b is any measurable subset of Ω where ρ is equal to zero or one a.e. There are no proven convergence results for the finite element approximation of the pressure, p . Since there can be multiple local minimizers, it is unclear which solution the sequence of finite element solutions is converging to. In this section, we consider any suitable conforming mixed finite element space such that the velocity and pressure spaces are inf–sup stable. We prove that, for every isolated minimizer of (BP), there exist sequences of finite element solutions to the discretized first-order optimality conditions that strongly converge to the minimizer as the mesh size tends to zero. More specifically, we show that, for each isolated minimizer of the infinite-dimensional problem, there exist (possibly different) sequences of finite element solutions $(\mathbf{u}_h, \rho_h, p_h)$ that converge to it strongly in $H^1(\Omega)^d \times L^s(\Omega) \times L^2(\Omega)$ as $h \rightarrow 0$, where $s \in [1, \infty)$. We emphasize that the results hold in the case where the local minima are isolated.

Consider the conforming finite element spaces $X_h \subset H^1(\Omega)^d$, $C_{\gamma,h} \subset C_\gamma$, and $M_h \subset L_0^2(\Omega)$. Let $X_{0,h} := \{\mathbf{v}_h \in X_h : \mathbf{v}_h|_{\partial\Omega} = \mathbf{0}\}$.

In general, it will not be possible to represent the boundary data \mathbf{g} exactly in the velocity finite element space. Hence, for each h , we instead consider boundary data \mathbf{g}_h (which can be represented) and assume that

$$(F1) \quad \mathbf{g}_h \rightarrow \mathbf{g} \text{ strongly in } H^{1/2}(\partial\Omega)^d.$$

We now define the space $X_{\mathbf{g}_h, h} := \{\mathbf{v}_h \in X_h : \mathbf{v}_h|_{\partial\Omega} = \mathbf{g}_h\}$. We will also assume that:

$$(F2) \quad X_{0, h} \text{ and } M_h \text{ satisfy the following inf-sup condition for some } c_b > 0, \text{ independent of } h,$$

$$c_b \leq \inf_{q_h \in M_h \setminus \{0\}} \sup_{\mathbf{v}_h \in X_{0, h} \setminus \{0\}} \frac{b(\mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{H^1(\Omega)} \|q_h\|_{L^2(\Omega)}}. \quad (4.1)$$

$$(F3) \quad \text{The finite element spaces are dense in their respective function spaces, i.e., for any } (\mathbf{v}, \eta, q) \in H^1(\Omega)^d \times C_\gamma \times L_0^2(\Omega),$$

$$\begin{aligned} \lim_{h \rightarrow 0} \inf_{\mathbf{w}_h \in X_h} \|\mathbf{v} - \mathbf{w}_h\|_{H^1(\Omega)} &= \lim_{h \rightarrow 0} \inf_{\zeta_h \in C_{\gamma, h}} \|\eta - \zeta_h\|_{L^2(\Omega)} \\ &= \lim_{h \rightarrow 0} \inf_{r_h \in M_h} \|q - r_h\|_{L^2(\Omega)} = 0. \end{aligned}$$

Theorem 4 (Convergence of the Finite Element Method). *Let $\Omega \subset \mathbb{R}^d$ be a polygonal domain in two dimensions or a polyhedral Lipschitz domain in three dimensions. Suppose that (A1)–(A5) hold and there exists an isolated local minimizer $(\mathbf{u}, \rho) \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma$ of (BP). Moreover, assume that, for $\theta > 0$, \bar{U}_θ is the subset of Ω where $|\mathbf{u}|^2 \geq \theta$ a.e. in \bar{U}_θ and suppose that there exists a $\theta' > 0$ such that \bar{U}_θ is closed and has non-empty interior U_θ for all $\theta \leq \theta'$. Let p denote the unique Lagrange multiplier associated with (\mathbf{u}, ρ) such that (\mathbf{u}, ρ, p) satisfy the first-order optimality conditions (FOC1)–(FOC3).*

Consider the conforming finite element spaces $X_h \subset H^1(\Omega)^d$, $C_{\gamma, h} \subset C_\gamma$, and $M_h \subset L_0^2(\Omega)$ and suppose that the assumptions (F1)–(F3) hold.

Then, there exists an $\bar{h} > 0$ such that, for $\bar{h} \geq h \rightarrow 0$, there is a sequence of solutions $(\mathbf{u}_h, \rho_h, p_h) \in X_{\mathbf{g}_h, h} \times C_{\gamma, h} \times M_h$ to the following discretized first-order optimality conditions

$$a_{\rho_h}(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = l_f(\mathbf{v}_h) \quad \text{for all } \mathbf{v}_h \in X_{0, h}, \quad (\text{FOC1}_h)$$

$$b(\mathbf{u}_h, q_h) = 0 \quad \text{for all } q_h \in M_h, \quad (\text{FOC2}_h)$$

$$c_{\mathbf{u}_h}(\rho_h, \eta_h - \rho_h) \geq 0 \quad \text{for all } \eta_h \in C_{\gamma, h}, \quad (\text{FOC3}_h)$$

such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$, $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$, $s \in [1, \infty)$, and $p_h \rightarrow p$ strongly in $L^2(\Omega)$ as $h \rightarrow 0$.

In Proposition 4, by fixing a ball around an isolated local minimizer, we show that finite element minimizers of a modified optimization problem converge weakly in $H^1(\Omega)^d \times L^2(\Omega)$ to the isolated minimizer of the infinite-dimensional problem. From this we deduce that there exists a subsequence of finite element solutions (\mathbf{u}_h) that converges strongly to the isolated minimizer of the infinite-dimensional problem in $L^2(\Omega)^d$. We then strengthen the convergence of ρ_h to strong convergence in $L^s(\Omega)$, $s \in [1, \infty)$, in Proposition 5 and strengthen the convergence of \mathbf{u}_h to strong convergence in $H^1(\Omega)^d$ in Proposition 6. In Proposition 7, we prove that there exists an $\bar{h} > 0$ such that there is a subsequence, $\bar{h} > h \rightarrow 0$, of strongly converging finite element solutions that also satisfy discretized first-order optimality conditions. Finally, in Proposition 8, we show that the Lagrange multiplier, $p_h \in M_h$, that satisfies the discretized first-order optimality conditions, converges strongly in $L^2(\Omega)$ to the Lagrange multiplier for the infinite-dimensional problem.

We now fix an isolated minimizer (\mathbf{u}, ρ) of (BP). We define the radius of the basin of attraction as the largest value r such that (\mathbf{u}, ρ) is the unique local minimizer in $B_{r, H^1(\Omega) \times L^2(\Omega)}(\mathbf{u}, \rho) \cap (H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma)$, where

$$\begin{aligned} B_{r, H^1(\Omega) \times L^2(\Omega)}(\mathbf{u}, \rho) \\ := \{\mathbf{v} \in H^1(\Omega)^d, \eta \in C_\gamma : \|\mathbf{u} - \mathbf{v}\|_{H^1(\Omega)} + \|\rho - \eta\|_{L^2(\Omega)} \leq r\}. \end{aligned} \quad (4.2)$$

We also define $B_{r, H^1(\Omega)}(\mathbf{u})$ and $B_{r, L^2(\Omega)}(\rho)$ by

$$B_{r, H^1(\Omega)}(\mathbf{u}) := \{\mathbf{v} \in H^1(\Omega)^d : \|\mathbf{u} - \mathbf{v}\|_{H^1(\Omega)} \leq r\}, \quad (4.3)$$

$$B_{r, L^2(\Omega)}(\rho) := \{\eta \in C_\gamma : \|\rho - \eta\|_{L^2(\Omega)} \leq r\}. \quad (4.4)$$

We note that

$$\begin{aligned} (H_{\mathbf{g}, \text{div}}^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho)) \\ \subset B_{r, H^1(\Omega) \times L^2(\Omega)}(\mathbf{u}, \rho) \cap (H_{\mathbf{g}, \text{div}}^1(\Omega)^d \times C_\gamma) \end{aligned}$$

and hence (\mathbf{u}, ρ) is also the unique minimizer in $(H_{\mathbf{g}, \text{div}}^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho))$.

Moreover, we define the spaces $V_{\mathbf{g}_h, h}$ and $V_{0, h}$ by

$$V_{\mathbf{g}_h, h} := \{\mathbf{v}_h \in X_{\mathbf{g}_h, h} : b(\mathbf{v}_h, q_h) = 0 \text{ for all } q_h \in M_h\},$$

$$V_{0, h} := \{\mathbf{v}_h \in X_{0, h} : b(\mathbf{v}_h, q_h) = 0 \text{ for all } q_h \in M_h\}.$$

Remark 7. In the context of the material distribution, it might be a more natural choice to assume that ρ is isolated with respect to the L^∞ -norm. Assuming ρ is isolated with respect to the L^2 -norm is a stronger isolation assumption, as it cannot be guaranteed that if $\eta \in C_\gamma$ lives in an isolated neighborhood with respect to the L^2 -norm, $\eta \in B_{r,L^2(\Omega)}(\rho)$, then, there exists an $r_* > 0$, such that $\eta \in B_{r_*,L^\infty(\Omega)}(\rho)$ where $B_{r_*,L^\infty(\Omega)}(\rho)$ is an isolated neighborhood with respect to the L^∞ -norm. We make this stronger isolation assumption as simple and continuous functions are not dense in $L^\infty(\Omega)$, but are dense in $L^2(\Omega)$. This has implications in the assumption (F3) and, subsequently, in the remaining results. However, as far as we are aware, the L^2 -isolation assumption is valid for all practical problems found in the literature, in particular it holds for both examples found in Section 5.

Remark 8. We note that balls centered at $\rho \in C_\gamma$ are equivalent if measured against any L^s -norm for $s \in [1, \infty)$. More precisely, for any $s, q \in [1, \infty)$, if $\eta \in C_\gamma \cap B_{r,L^q(\Omega)}(\rho)$, there exists an $r_* > 0$, depending on $s \in [1, \infty)$, such that $\eta \in C_\gamma \cap B_{r_*,L^s(\Omega)}(\rho)$. Hence, the assumption that the material distribution is isolated with respect to L^2 -norm is not a stronger assumption than being isolated with respect to the L^s -norm provided $s \in [1, \infty)$.

In Propositions 4–7 and Corollary 1, we fix an isolated minimizer (\mathbf{u}, ρ) of (BP) and suppose that the conditions of Theorem 4 hold.

Proposition 4 (Weak Convergence of (\mathbf{u}_h, ρ_h) in $H^1(\Omega)^d \times L^2(\Omega)$). Consider the finite-dimensional optimization problem: find (\mathbf{u}_h, ρ_h) that minimizes

$$\min_{(\mathbf{v}_h, \eta_h) \in (V_{\mathbf{g}_h, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho))} J(\mathbf{v}_h, \eta_h). \quad (\text{BP}_h)$$

Then, a minimizer (\mathbf{u}_h, ρ_h) of (BP_h) exists and there exist subsequences (up to relabeling) such that

$$\mathbf{u}_h \rightharpoonup \mathbf{u} \text{ weakly in } H^1(\Omega)^d, \quad (4.5)$$

$$\mathbf{u}_h \rightarrow \mathbf{u} \text{ strongly in } L^2(\Omega)^d, \quad (4.6)$$

$$\rho_h \xrightarrow{*} \rho \text{ weakly-}^* \text{ in } L^\infty(\Omega), \quad (4.7)$$

$$\rho_h \rightharpoonup \rho \text{ weakly in } L^s(\Omega), \quad s \in [1, \infty). \quad (4.8)$$

Proof. The functional J is continuous and $(V_{\mathbf{g}_h, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho))$ is a finite-dimensional, closed and bounded set. Moreover, for sufficiently small h it is non-empty. Therefore, it is sequentially compact by the Heine–Borel theorem. Hence J obtains its infimum in $(V_{\mathbf{g}_h, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho))$ and therefore, a minimizer (\mathbf{u}_h, ρ_h) exists.

By a corollary of Kakutani's Theorem, if a Banach space is reflexive then every norm-closed, bounded and convex subset of the Banach space is weakly compact and thus, by the Eberlein–Šmulian theorem, sequentially weakly compact. It can be checked that $H^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ and $C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho)$ are norm-closed, bounded and convex subsets of the reflexive Banach spaces $H^1(\Omega)^d$ and $L^2(\Omega)$, respectively. Therefore, $H^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ is weakly sequentially compact in $H^1(\Omega)^d$ and $C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho)$ is weakly sequentially compact in $L^2(\Omega)$.

Hence we can extract subsequences, (\mathbf{u}_h) and (ρ_h) of the sequence generated by the global minimizers of (BP_h) such that

$$\mathbf{u}_h \rightharpoonup \hat{\mathbf{u}} \in H^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u}) \text{ weakly in } H^1(\Omega)^d, \quad (4.9)$$

$$\rho_h \rightharpoonup \hat{\rho} \in C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho) \text{ weakly in } L^2(\Omega). \quad (4.10)$$

By assumption (F3), there exists a sequence of finite element functions $\tilde{\rho}_h \in C_{\gamma, h}$ that strongly converges to ρ in $L^2(\Omega)$. Moreover let $\tilde{\mathbf{u}}_h \in V_{\mathbf{g}_h, h}$ be a finite element function taken from the sequence of finite element functions that satisfy $\tilde{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Such a sequence is shown to exist in [6, Lemma 3.1].

We now wish to identify the limits $\hat{\mathbf{u}}$ and $\hat{\rho}$. Consider the following bound:

$$\begin{aligned} & |J(\tilde{\mathbf{u}}_h, \tilde{\rho}_h) - J(\mathbf{u}, \rho)| \\ & \leq \int_{\Omega} |(\alpha(\rho) - \alpha(\tilde{\rho}_h))|\mathbf{u}|^2| + |\alpha(\tilde{\rho}_h)(|\mathbf{u}|^2 - |\tilde{\mathbf{u}}_h|^2)| \, dx \\ & \quad + \int_{\Omega} \nu ||\nabla \mathbf{u}|^2 - |\nabla \tilde{\mathbf{u}}_h|^2| + 2|\mathbf{f} \cdot (\mathbf{u} - \tilde{\mathbf{u}}_h)| \, dx \\ & \leq L_\alpha \|\mathbf{u}\|_{L^4(\Omega)}^2 \|\tilde{\rho}_h - \rho\|_{L^2(\Omega)} \\ & \quad + \tilde{\alpha} \|\tilde{\mathbf{u}}_h - \mathbf{u}\|_{L^2(\Omega)} (\|\tilde{\mathbf{u}}_h - \mathbf{u}\|_{L^2(\Omega)} + 2\|\mathbf{u}\|_{L^2(\Omega)}) \\ & \quad + \nu \|\tilde{\mathbf{u}}_h - \mathbf{u}\|_{H^1(\Omega)} (\|\tilde{\mathbf{u}}_h - \mathbf{u}\|_{H^1(\Omega)} + 2\|\mathbf{u}\|_{H^1(\Omega)}) \\ & \quad + 2\|\mathbf{f}\|_{L^2(\Omega)} \|\tilde{\mathbf{u}}_h - \mathbf{u}\|_{L^2(\Omega)}, \end{aligned} \quad (4.11)$$

where L_α denotes the Lipschitz constant for α . From (4.11) we see that

$$J(\tilde{\mathbf{u}}_h, \tilde{\rho}_h) \rightarrow J(\mathbf{u}, \rho) \text{ as } h \rightarrow 0.$$

Furthermore, for sufficiently small $h > 0$ we note that

$$(\tilde{\mathbf{u}}_h, \tilde{\rho}_h) \in (V_{\mathbf{g}_h, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho)).$$

Therefore,

$$J(\mathbf{u}_h, \rho_h) \leq J(\tilde{\mathbf{u}}_h, \tilde{\rho}_h). \quad (4.12)$$

By taking the limit as $h \rightarrow 0$ and utilizing the strong convergence of $\tilde{\mathbf{u}}_h$ and $\tilde{\rho}_h$ to \mathbf{u} and ρ , respectively, we see that

$$\lim_{h \rightarrow 0} J(\mathbf{u}_h, \rho_h) \leq J(\mathbf{u}, \rho). \quad (4.13)$$

(F1) implies that

$$\mathbf{u}_h|_{\partial\Omega} = \mathbf{g}_h \rightarrow \mathbf{g} \text{ strongly in } H^{1/2}(\partial\Omega)^d. \quad (4.14)$$

By assumption (F3), for every $q \in L_0^2(\Omega)$, there exists a sequence of $\tilde{q}_h \in M_h$ such that $\tilde{q}_h \rightarrow q$ strongly in $L^2(\Omega)$. Since $\mathbf{u}_h \rightharpoonup \hat{\mathbf{u}}$ weakly in $H^1(\Omega)^d$ and $\mathbf{u}_h \in V_{\mathbf{g}_h, h}$, we see that

$$b(\hat{\mathbf{u}}, q) = \lim_{h \rightarrow 0} b(\mathbf{u}_h, \tilde{q}_h) + \lim_{h \rightarrow 0} b(\mathbf{u}_h, q - \tilde{q}_h) = 0 \text{ for all } q \in L_0^2(\Omega). \quad (4.15)$$

Hence $\hat{\mathbf{u}}$ is pointwise divergence-free and together with (4.14), we deduce that $\hat{\mathbf{u}} \in H_{\mathbf{g}, \text{div}}^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$. By construction $\hat{\rho} \in C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho)$.

With a small modification to the proof found in [6, Th. 3.1], we note that J is weakly lower semicontinuous on $H^1(\Omega)^d \times L^2(\Omega)$. Therefore,

$$J(\hat{\mathbf{u}}, \hat{\rho}) \leq \liminf_{h \rightarrow 0} J(\mathbf{u}_h, \rho_h). \quad (4.16)$$

We note that (\mathbf{u}, ρ) is the unique minimizer of $(H_{\mathbf{g}, \text{div}}^1(\Omega)^d \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (C_\gamma \cap B_{r/2, L^2(\Omega)}(\rho))$, which implies that $J(\mathbf{u}, \rho) \leq J(\hat{\mathbf{u}}, \hat{\rho})$. Hence, from (4.13) and (4.16), it follows that

$$J(\hat{\mathbf{u}}, \hat{\rho}) = J(\mathbf{u}, \rho). \quad (4.17)$$

Since (\mathbf{u}, ρ) is the unique minimizer in the spaces we consider, we can identify the limits $\hat{\mathbf{u}}$ and $\hat{\rho}$ as \mathbf{u} and ρ , respectively, and state that $\mathbf{u}_h \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$ and $\rho_h \rightharpoonup \rho$ weakly in $L^2(\Omega)$. By the Rellich–Kondrachov theorem, we can extract a further subsequence such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $L^2(\Omega)^d$.

We note that by the Banach–Alaoglu theorem, the closed unit ball of the dual space of a normed vector space, (for example $L^1(\Omega)$), is compact in the weak-* topology. Hence we can also find a subsequence such that $\rho_h \xrightarrow{*} \hat{\rho} \in C_\gamma \cap \{\eta : \|\rho - \eta\|_{L^\infty(\Omega)} \leq r/2\}$ weakly-* in $L^\infty(\Omega)$. By the uniqueness of the weak limit, we can identify $\hat{\rho} = \rho$ a.e. in Ω and, thus, we deduce that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$, i.e. $\int_\Omega \rho_h \eta \, dx \rightarrow \int_\Omega \rho \eta \, dx$ for any $\eta \in L^1(\Omega)$. Since $L^s(\Omega) \subset L^1(\Omega)$, for any $s > 1$, we note that $\int_\Omega \rho_h \eta \, dx \rightarrow \int_\Omega \rho \eta \, dx$ for any $\eta \in L^s(\Omega)$, $s \geq 1$. Hence, by definition, $\rho_h \rightharpoonup \rho$ weakly in $L^s(\Omega)$ for all $s \in [1, \infty)$. \square

Corollary 1 (Strong Convergence of ρ_h in $L^s(\Omega_b)$). *Let Ω_b be any measurable subset of Ω of positive measure on which ρ is equal to zero or one a.e. (if such a set exists). Then, there exists a sequence of finite element minimizers, ρ_h , of (BP_h) that converge strongly in $L^s(\Omega_b)$ to the isolated local minimizer of (BP), where $s \in [1, \infty)$.*

Proof. We have shown that there exists a sequence of finite element minimizers (\mathbf{u}_h, ρ_h) of (BP_h) that converge to the isolated local minimizer (\mathbf{u}, ρ) . In particular $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$ and $\rho_h \rightharpoonup \rho$ weakly in $L^s(\Omega)$ for all $s \in [1, \infty)$. The result is then deduced by following the proof of Corollary 3.2 in [27]. \square

Proposition 5 (Strong Convergence of ρ_h in $L^s(\Omega)$, $s \in [1, \infty)$). *There exists a subsequence of minimizers, (ρ_h) , of (BP_h) such that*

$$\rho_h \rightarrow \rho \text{ strongly in } L^s(\Omega), \quad s \in [1, \infty). \quad (4.18)$$

Proof. We note that $C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho)$ is a convex set, and hence for any $\eta_h \in C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho)$, $t \in [0, 1]$, we have that $\rho_h + t(\eta_h - \rho_h) \in C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho)$. Since (\mathbf{u}_h, ρ_h) is a minimizer of (BP_h), by the arguments used in Proposition 2 we deduce that

$$\int_\Omega \alpha'(\rho_h) |\mathbf{u}_h|^2 (\eta_h - \rho_h) \, dx \geq 0 \text{ for all } \eta_h \in C_{\gamma, h} \cap B_{r/2, L^2(\Omega)}(\rho). \quad (4.19)$$

Hence, (FOC3) and (4.19) imply that for all $\eta \in C_\gamma$ and $\eta_h \in C_{\gamma,h} \cap B_{r/2,L^2(\Omega)}(\rho)$ we have that

$$\int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 \rho \, dx \leq \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 \eta \, dx, \quad (4.20)$$

$$\int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 \rho_h \, dx \leq \int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 \eta_h \, dx. \quad (4.21)$$

By subtracting $\int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 \rho_h \, dx$ from (4.20) and $\int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 \rho \, dx$ from (4.21), we see that

$$\int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\rho - \rho_h) \, dx \leq \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\eta - \rho_h) \, dx, \quad (4.22)$$

$$\int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 (\rho_h - \rho) \, dx \leq \int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 (\eta_h - \rho) \, dx. \quad (4.23)$$

Summing (4.22) and (4.23) and rearranging the left-hand side, we see that

$$\begin{aligned} & \int_{\Omega} (\alpha'(\rho) - \alpha'(\rho_h)) |\mathbf{u}|^2 (\rho - \rho_h) \, dx + \int_{\Omega} \alpha'(\rho_h) (|\mathbf{u}|^2 - |\mathbf{u}_h|^2) (\rho - \rho_h) \, dx \\ & \leq \int_{\Omega} \alpha'(\rho) |\mathbf{u}|^2 (\eta - \rho_h) \, dx + \int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 (\eta_h - \rho) \, dx. \end{aligned} \quad (4.24)$$

By fixing $\eta = \rho_h \in C_\gamma$ and subtracting the second term on the left-hand side of (4.24) from both sides we deduce that

$$\begin{aligned} & \int_{\Omega} (\alpha'(\rho) - \alpha'(\rho_h)) |\mathbf{u}|^2 (\rho - \rho_h) \, dx \\ & \leq \int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 (\eta_h - \rho) \, dx + \int_{\Omega} \alpha'(\rho_h) (|\mathbf{u}_h|^2 - |\mathbf{u}|^2) (\rho - \rho_h) \, dx. \end{aligned} \quad (4.25)$$

By an application of the mean value theorem, we note that there exists a $c \in (0, 1)$ such that

$$\begin{aligned} & \int_{\Omega} (\alpha'(\rho) - \alpha'(\rho_h)) |\mathbf{u}|^2 (\rho - \rho_h) \, dx \\ & = \int_{\Omega} \alpha''(\rho_h + c(\rho - \rho_h)) |\mathbf{u}|^2 (\rho - \rho_h)^2 \, dx. \end{aligned} \quad (4.26)$$

By (A5) and the definition of U_θ we bound (4.26) from below:

$$\begin{aligned} & \int_{\Omega} \alpha''(\rho_h + c(\rho - \rho_h)) |\mathbf{u}|^2 (\rho - \rho_h)^2 \, dx \\ & \geq \int_{U_\theta} \alpha''(\rho_h + c(\rho - \rho_h)) |\mathbf{u}|^2 (\rho - \rho_h)^2 \, dx \geq \alpha''_{\min} \theta \|\rho - \rho_h\|_{L^2(U_\theta)}^2. \end{aligned} \quad (4.27)$$

Now we bound the right-hand side of (4.25) as follows,

$$\begin{aligned} & \int_{\Omega} \alpha'(\rho_h) |\mathbf{u}_h|^2 (\eta_h - \rho) \, dx + \int_{\Omega} \alpha'(\rho_h) (|\mathbf{u}_h|^2 - |\mathbf{u}|^2) (\rho - \rho_h) \, dx \\ & \leq 2\alpha'_{\max} (\|\mathbf{u}\|_{L^4(\Omega)}^2 + \|\mathbf{u} - \mathbf{u}_h\|_{L^4(\Omega)}^2) \|\rho - \eta_h\|_{L^2(\Omega)} \\ & \quad + \alpha'_{\max} \|\rho - \rho_h\|_{L^q(\Omega)} \|\mathbf{u} + \mathbf{u}_h\|_{L^{q'}(\Omega)} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}, \end{aligned} \quad (4.28)$$

where $2 < q' < \infty$ in two dimensions, $2 < q' \leq 6$ in three dimensions, and $q = 2q'/(q' - 2)$. We note that

$$\begin{aligned} \|\mathbf{u} + \mathbf{u}_h\|_{L^{q'}(\Omega)} & \leq \|\mathbf{u}\|_{L^{q'}(\Omega)} + \|\mathbf{u}_h\|_{L^{q'}(\Omega)} \\ & \leq \|\mathbf{u}\|_{H^1(\Omega)} + \|\mathbf{u}_h\|_{H^1(\Omega)} \leq \hat{C} < \infty, \end{aligned} \quad (4.29)$$

where the second inequality holds thanks to the Sobolev embedding theorem. Combining (4.25)–(4.29) we see that

$$\|\rho - \rho_h\|_{L^2(U_\theta)}^2 \leq C (\|\rho - \eta_h\|_{L^2(\Omega)} + \|\rho - \rho_h\|_{L^q(\Omega)} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}), \quad (4.30)$$

where $C = C(\alpha'_{\max}, \alpha''_{\min}, \theta, \|\mathbf{u}\|_{L^4(\Omega)}, \hat{C})$. By assumption (F3), there exists a sequence of finite element functions $\tilde{\rho}_h \in C_{\gamma,h}$ such that $\tilde{\rho}_h \rightarrow \rho$ strongly in $L^2(\Omega)$. Thanks to the strong convergence, we note that for sufficiently small h , $\tilde{\rho}_h \in C_{\gamma,h} \cap B_{r/2,L^2(\Omega)}(\rho)$. Hence we can fix $\eta_h = \tilde{\rho}_h$. By Proposition 4, we know that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $L^2(\Omega)^d$ and since $\rho \in C_\gamma$, $\rho_h \in C_{\gamma,h} \subset C_\gamma$, then $\|\rho - \rho_h\|_{L^q(\Omega)} \leq |\Omega|^{1/q} \|\rho - \rho_h\|_{L^\infty(\Omega)} \leq |\Omega|^{1/q}$. Therefore, the right-hand side of (4.30) tends to zero as $h \rightarrow 0$. Hence, we deduce that

$$\rho_h \rightarrow \rho \text{ strongly in } L^2(U_\theta), \quad \theta > 0. \quad (4.31)$$

Now we note that

$$\|\rho - \rho_h\|_{L^2(\Omega)} = \|\rho - \rho_h\|_{L^2(U_\theta)} + \|\rho - \rho_h\|_{L^2(U \setminus U_\theta)} + \|\rho - \rho_h\|_{L^2(\Omega \setminus U)}. \quad (4.32)$$

If $U \setminus U_\theta$ or $\Omega \setminus U$ are empty, we neglect the corresponding term in (4.32) with no loss of generality. Suppose $\Omega \setminus U$ is non-empty. By definition of U , $\mathbf{u} = \mathbf{0}$ a.e. in $\Omega \setminus U$. By Proposition 3, this implies that $\rho = 0$ a.e. in $\Omega \setminus U$. Therefore, Corollary 1 implies that

$$\rho_h \rightarrow \rho \text{ strongly in } L^2(\Omega \setminus U). \quad (4.33)$$

Suppose $U \setminus U_\theta$ is non-empty. Since, $\rho, \rho_h \in C_Y$ we see that

$$\|\rho - \rho_h\|_{L^2(U \setminus U_\theta)} \leq |U \setminus U_\theta|^{1/2} \rightarrow 0 \text{ as } \theta \rightarrow 0. \quad (4.34)$$

Therefore, by first taking the limit as $h \rightarrow 0$ and then by taking the limit as $\theta \rightarrow 0$, (4.31)–(4.34) imply that $\rho_h \rightarrow \rho$ strongly in $L^2(\Omega)$.

Since $\|\rho - \rho_h\|_{L^1(\Omega)} \leq |\Omega|^{1/2} \|\rho - \rho_h\|_{L^2(\Omega)}$, we see that $\rho_h \rightarrow \rho$ strongly in $L^1(\Omega)$. Hence, for any $s \in [1, \infty)$,

$$\int_{\Omega} |\rho - \rho_h|^s dx = \int_{\Omega} |\rho - \rho_h|^{s-1} |\rho - \rho_h| dx \leq 1^{s-1} \|\rho - \rho_h\|_{L^1(\Omega)}, \quad (4.35)$$

which implies that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$. \square

Proposition 6 (Strong Convergence of \mathbf{u}_h in $H^1(\Omega)^d$). *There exists a subsequence of minimizers, (\mathbf{u}_h) , of (BP_h) such that*

$$\mathbf{u}_h \rightarrow \mathbf{u} \text{ strongly in } H^1(\Omega)^d. \quad (4.36)$$

Proof. We note that the set $W_h := V_{\mathbf{g}_h, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ is convex. Hence, by following the same arguments for deriving the variational inequality on ρ as in Proposition 2, it can be shown that minimizers of (BP_h) satisfy the variational inequality

$$a_{\rho_h}(\mathbf{u}_h, \mathbf{w}_h - \mathbf{u}_h) - l_f(\mathbf{w}_h - \mathbf{u}_h) \geq 0 \text{ for all } \mathbf{w}_h \in W_h. \quad (4.37)$$

We note that in the next proposition (Proposition 7), we will show that (4.37) can be strengthened to an equality. However, this result is not currently available at this point and an equality does not follow from the arguments in Proposition 2. We note though that an inequality is sufficient for the subsequent arguments. From Proposition 2 we deduce that, for all $\mathbf{w}_h \in W_h$,

$$a_{\rho}(\mathbf{u}, \mathbf{w}_h - \mathbf{u}_h) + b(\mathbf{w}_h - \mathbf{u}_h, p) = l_f(\mathbf{w}_h - \mathbf{u}_h).$$

Hence

$$a_{\rho_h}(\mathbf{u}_h, \mathbf{u}_h - \mathbf{w}_h) \leq a_{\rho}(\mathbf{u}, \mathbf{u}_h - \mathbf{w}_h) + b(\mathbf{u}_h - \mathbf{w}_h, p). \quad (4.38)$$

By subtracting $a_{\rho_h}(\mathbf{w}_h, \mathbf{u}_h - \mathbf{w}_h)$ from both sides we see that

$$\begin{aligned} a_{\rho_h}(\mathbf{u}_h - \mathbf{w}_h, \mathbf{u}_h - \mathbf{w}_h) \\ \leq a_{\rho}(\mathbf{u}, \mathbf{u}_h - \mathbf{w}_h) - a_{\rho_h}(\mathbf{w}_h, \mathbf{u}_h - \mathbf{w}_h) + b(\mathbf{u}_h - \mathbf{w}_h, p - q_h). \end{aligned}$$

We note that a_{ρ_h} is coercive with constant $c_a = \nu/(c_p^2 + 1)$, and b is bounded with constant C_b . Hence,

$$\begin{aligned} \|\mathbf{u}_h - \mathbf{w}_h\|_{H^1(\Omega)}^2 &\leq \frac{1}{c_a} a_{\rho_h}(\mathbf{u}_h - \mathbf{w}_h, \mathbf{u}_h - \mathbf{w}_h) \\ &\leq \frac{1}{c_a} (a_{\rho}(\mathbf{u}, \mathbf{u}_h - \mathbf{w}_h) - a_{\rho_h}(\mathbf{w}_h, \mathbf{u}_h - \mathbf{w}_h) + b(\mathbf{u}_h - \mathbf{w}_h, p - q_h)) \\ &= \frac{1}{c_a} \left(\int_{\Omega} \alpha(\rho_h)(\mathbf{u} - \mathbf{w}_h) \cdot (\mathbf{u}_h - \mathbf{w}_h) + (\alpha(\rho) - \alpha(\rho_h)) \mathbf{u} \cdot (\mathbf{u}_h - \mathbf{w}_h) dx \right. \\ &\quad \left. + \int_{\Omega} \nu \nabla(\mathbf{u} - \mathbf{w}_h) : \nabla(\mathbf{u}_h - \mathbf{w}_h) dx + b(\mathbf{u}_h - \mathbf{w}_h, p - q_h) \right) \\ &\leq \frac{1}{c_a} \bar{\alpha} \|\mathbf{u} - \mathbf{w}_h\|_{L^2(\Omega)} \|\mathbf{u}_h - \mathbf{w}_h\|_{L^2(\Omega)} \\ &\quad + \frac{1}{c_a} \|(\alpha(\rho) - \alpha(\rho_h)) \mathbf{u}\|_{L^2(\Omega)} \|\mathbf{u}_h - \mathbf{w}_h\|_{L^2(\Omega)} \\ &\quad + \frac{\nu}{c_a} \|\mathbf{u} - \mathbf{w}_h\|_{H^1(\Omega)} \|\mathbf{u}_h - \mathbf{w}_h\|_{H^1(\Omega)} + \frac{C_b}{c_a} \|\mathbf{u}_h - \mathbf{w}_h\|_{H^1(\Omega)} \|p - q_h\|_{L^2(\Omega)}. \end{aligned}$$

Hence,

$$\|\mathbf{u}_h - \mathbf{w}_h\|_{H^1(\Omega)} \leq C (\|\mathbf{u} - \mathbf{w}_h\|_{H^1(\Omega)} + \|(\alpha(\rho) - \alpha(\rho_h)) \mathbf{u}\|_{L^2(\Omega)} + \|p - q_h\|_{L^2(\Omega)}),$$

where $C = C(\bar{\alpha}, \nu, c_a, C_b)$ is a constant. This implies that, for all $\mathbf{w}_h \in W_h$,

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} &\leq C' \left(\|\mathbf{u} - \mathbf{w}_h\|_{H^1(\Omega)} + \|(\alpha(\rho) - \alpha(\rho_h))\mathbf{u}\|_{L^2(\Omega)} + \|p - q_h\|_{L^2(\Omega)} \right). \end{aligned} \quad (4.39)$$

where $C' = C'(\bar{\alpha}, \nu, c_a, C_b, L_\alpha)$. For sufficiently small h , we note that $\tilde{\mathbf{u}}_h \in W_h$ (where $\tilde{\mathbf{u}}_h$ is defined in the proof of Proposition 4) and $\tilde{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Moreover by assumption (F3), there exists a sequence of finite element functions $\tilde{p}_h \in M_h$ that converges to p strongly in $L^2(\Omega)$. Suppose $\mathbf{w}_h = \tilde{\mathbf{u}}_h$ and $q_h = \tilde{p}_h$. From Proposition 5, we know that there exists a subsequence (not indicated) such that $\rho_h \rightarrow \rho$ strongly in $L^4(\Omega)$. We now observe that

$$\|(\alpha(\rho) - \alpha(\rho_h))\mathbf{u}\|_{L^2(\Omega)} \leq L_\alpha \|\rho - \rho_h\|_{L^4(\Omega)} \|\mathbf{u}\|_{L^4(\Omega)} \quad (4.40)$$

where L_α is the Lipschitz constant for α . Hence by taking the limit as $h \rightarrow 0$, we deduce that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. \square

In the following proposition, we show that (up to a subsequence) minimizers of (BP_h) also satisfy the first-order optimality conditions that are the finite-dimensional analogue of the first-order optimality conditions associated with (BP). This allows us to consider the finite-dimensional optimization problem over the whole set $V_{g_h,h} \times C_{\gamma,h}$, rather than the restricted set $(V_{g_h,h} \cap B_{r/2,H^1(\Omega)}(\mathbf{u})) \times (C_{\gamma,h} \cap B_{r/2,L^2(\Omega)}(\rho))$.

Proposition 7 (Discretized First-Order Optimality Conditions). *There exists an $\bar{h} > 0$ such that for all $h < \bar{h}$, there exists a unique Lagrange multiplier $p_h \in M_h$ such that the functions (\mathbf{u}_h, ρ_h) that locally minimize (BP_h) satisfy the first-order optimality conditions (FOC1_h)–(FOC3_h)*

Proof. From Proposition 6, we know that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Hence by definition of strong convergence, there exists an $\bar{h} > 0$ such that, for all $h \leq \bar{h}$, $\|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} \leq r/4$. Therefore, for each $\mathbf{v}_h \in V_{0,h}$, if $|t| < r/(4\|\mathbf{v}_h\|_{H^1(\Omega)})$ then $\mathbf{u}_h + t\mathbf{v}_h \in V_{g_h,h} \cap B_{r/2,H^1(\Omega)}(\mathbf{u})$. Now we can follow the reasoning of the proof of Proposition 2 (adding the subscript h where necessary) to deduce the existence of a unique $p_h \in M_h$ such that (FOC1_h)–(FOC3_h) hold. \square

Proposition 8 (Strong Convergence of p_h in $L^2(\Omega)$). *There is a subsequence of the unique $p_h \in M_h$ defined in Proposition 7 that converges strongly in $L^2(\Omega)$ to the $p \in L_0^2(\Omega)$ that solves (FOC1)–(FOC3) for the given isolated local minimizer (\mathbf{u}, ρ) .*

Proof. The inf-sup condition for M_h and $X_{0,h}$ implies that, for any $q_h \in M_h$,

$$\begin{aligned} c_b \|q_h - p_h\|_{L^2(\Omega)} &\leq \sup_{\mathbf{w}_h \in X_{0,h} \setminus \{0\}} \frac{b(\mathbf{w}_h, q_h - p_h)}{\|\mathbf{w}_h\|_{H^1(\Omega)}} \\ &= \sup_{\mathbf{w}_h \in X_{0,h} \setminus \{0\}} \frac{b(\mathbf{w}_h, p - p_h) + b(\mathbf{w}_h, q_h - p)}{\|\mathbf{w}_h\|_{H^1(\Omega)}} \\ &\leq \sup_{\mathbf{w}_h \in X_{0,h} \setminus \{0\}} \frac{|b(\mathbf{w}_h, p - p_h)| + |b(\mathbf{w}_h, q_h - p)|}{\|\mathbf{w}_h\|_{H^1(\Omega)}} \\ &= \sup_{\mathbf{w}_h \in X_{0,h} \setminus \{0\}} \frac{|a_\rho(\mathbf{u}, \mathbf{w}_h) - a_{\rho_h}(\mathbf{u}_h, \mathbf{w}_h)| + |b(\mathbf{w}_h, q_h - p)|}{\|\mathbf{w}_h\|_{H^1(\Omega)}} \\ &\leq \|(\alpha(\rho) - \alpha(\rho_h))\mathbf{u}\|_{L^2(\Omega)} + (\bar{\alpha} + \nu) \|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} \\ &\quad + C_b \|p - q_h\|_{L^2(\Omega)}. \end{aligned}$$

Hence,

$$\|p - p_h\|_{L^2(\Omega)} \leq C \left(\|(\alpha(\rho) - \alpha(\rho_h))\mathbf{u}\|_{L^2(\Omega)} + \|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} + \|p - q_h\|_{L^2(\Omega)} \right),$$

where $C = C(c_b, C_b, \bar{\alpha}, L_\alpha, \nu)$. By assumption (F3), there exists a sequence of finite element functions, $\tilde{p}_h \in M_h$ that satisfies $\tilde{p}_h \rightarrow p$ strongly in $L^2(\Omega)$. Let $q_h = \tilde{p}_h$. We have already shown that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ in Proposition 6. Similarly, in the proof of Proposition 6 we also showed that $\|(\alpha(\rho) - \alpha(\rho_h))\mathbf{u}\|_{L^2(\Omega)} \rightarrow 0$. Hence we conclude that $p_h \rightarrow p$ strongly in $L^2(\Omega)$. \square

Proof of Theorem 4. Fix an isolated minimizer (\mathbf{u}, ρ) of (BP) and its unique associated Lagrange multiplier p . By the results of Propositions 4, 5, 6, and 7, there exists a mesh size \bar{h} such that for, $h < \bar{h}$, there exists a sequence of finite element solutions $(\mathbf{u}_h, \rho_h, p_h) \in V_{g_h,h} \times C_{\gamma,h} \times M_h$ satisfying (FOC1_h)–(FOC3_h) that converges to (\mathbf{u}, ρ, p) . By taking a subsequence if necessary (not indicated), Proposition 5, implies that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$, $s \in [1, \infty)$, Proposition 6 implies that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$, and Proposition 8 implies that $p_h \rightarrow p$ strongly in $L^2(\Omega)$. \square

5. Numerical results

The main goal of this section is to experimentally verify the existence of strongly converging sequences that were proven to exist in Section 4. In all examples the systems are discretized with the finite element method using FEniCS [28]. The computational domains are triangulated with simplices and we define the mesh size, h , as the maximum diameter of all the simplices in the triangulation. The solutions are computed using the deflated barrier method [29]. The deflated barrier method reformulates the discretized first-order optimality conditions (FOC1_h) – (FOC3_h) as a mixed complementarity problem. The mixed complementarity problem is then solved with a primal–dual active set strategy [30], a Newton-like solver that incorporates the box constraints on ρ_h . Typically, nonlinear convergence is difficult to achieve from a naïve initial guess. Hence, the functional J in (BP) is augmented with barrier-like terms. Continuation of the barrier parameter to zero then recovers the solution to the original first-order optimality conditions. A key feature of the deflated barrier method, as required in this work, is the ability to systematically discover multiple solutions of topology optimization problems from the same initial guess. This is achieved via the deflation technique [31,32], [29, Sec. 3.2]. Deflation prevents a Newton-like solver from converging to an already discovered solution by modifying the discretized first-order optimality conditions with a deflation operator. Deflation is extremely cheap to implement (effectively at the same cost as two inner products) and does not affect the conditioning of the linear systems that are solved during the run of the Newton-like solver. The resulting linear systems arising in the deflated barrier method are solved by a sparse LU factorization with MUMPS [33] and PETSc [34].

There are no known solutions for choices of the inverse permeability, α , used in practice. Hence, errors are measured with respect to a heavily-refined finite element solution, which is constructed as follows; first the finite element solutions are computed on a mesh with mesh size $h = 0.028$, using the deflated barrier method. Next, the mesh is adaptively refined three times in areas where the material distribution is between 1/10 and 9/10. Each time the mesh is refined, the coarse-mesh solution is interpolated onto the finer mesh as an initial guess and the first-order optimality conditions are re-solved using the deflated barrier method.

In principle, there can be infinitely many different subsequences of finite element solutions that strongly converge to the same minimizer of the infinite-dimensional problem at different convergence rates. Separate subsequences cause difficulties in the interpretation of the convergence plots as they present themselves as oscillations in the error. This is observed in practice and appears to be caused by at least the following two reasons:

- (P1) Multiple finite element solutions can exist on the same mesh that represent the same solution of the infinite-dimensional problem, e.g. Fig. 5;
- (P2) A fine mesh can align worse than a coarser mesh with the jumps in the material distribution that solves the infinite-dimensional problem.

Observation (P1) is not surprising in the context of nonlinear PDEs and nonconvex variational problems. In such cases, an additional selection mechanism is required in order to favor one particular solution over others coexisting on the same mesh. Selection mechanisms are problem-dependent. In the case of nonlinear hyperbolic conservation laws the entropy condition plays this role. In the present context, one might propose choosing the solution, minimizing the modified optimization problem (BP_h) , that attains the smallest objective functional value for J , within the basin of attraction of the isolated local minimizer. For sufficiently small h , a minimizer satisfying this selection mechanism must exist. However, it is not necessarily unique and numerically enforcing such a condition can be difficult. In order to promote convergence to the minimizer of (BP_h) with the smallest value J , we interpolate the heavily-refined finite element solutions onto coarser meshes as initial guesses for the deflated barrier method. This strategy was effective in practice. The effects of the second observation (P2) are harder to test. However, in Section 5.2, we attempt to minimize mesh bias by measuring errors on unstructured meshes.

Code availability: For reproducibility, an implementation of the deflated barrier method as well as scripts to generate the convergence plots and solutions can be found at <https://bitbucket.org/papadopoulos/deflatedbarrier/>. The version of the software used in this paper is archived on Zenodo [35].

5.1. Discontinuous-forcing

Consider the optimization problem (BP), with a homogeneous Dirichlet boundary condition on \mathbf{u} , $\Omega = (0, 1)^2$, volume fraction $\gamma = 1/3$, viscosity $\nu = 1$ and a forcing term given by

$$\mathbf{f}(x, y) = \begin{cases} (10, 0)^\top & \text{if } 3/10 < x < 7/10 \text{ and } 3/10 < y < 7/10, \\ (0, 0)^\top & \text{otherwise.} \end{cases} \quad (5.1)$$

The inverse permeability, α , is as given in (2.1), with $\bar{\alpha} = 2.5 \times 10^4$ and $q = 1/10$, which satisfies (A1)–(A5). Here q is a penalty parameter which controls the level of intermediate values (between zero or one) in the optimal design. Fig. 1 depicts the material distribution of three minimizers. One local minimizer is in the shape of a figure eight and the two \mathbb{Z}_2 symmetric global minimizers are in the shape of annuli. Since the domain is convex, $\mathbf{g} = \mathbf{0}$, and $\mathbf{f} \in L^2(\Omega)^d$, then, by the regularity results proven in the Appendix, $\mathbf{u} \in H^2(\Omega)^2$ and $p \in H^1(\Omega)$. The conditions of Theorem 3 hold and, therefore,

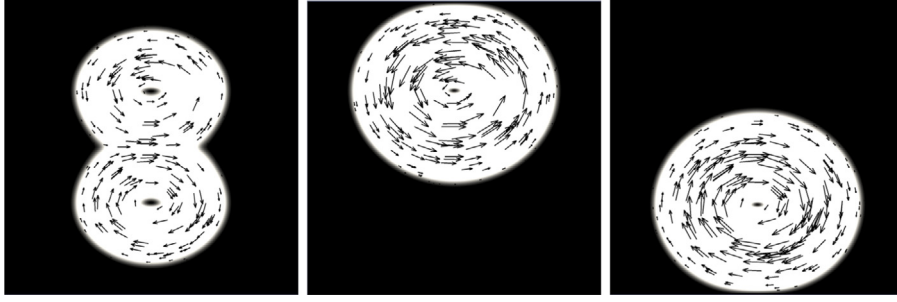


Fig. 1. The material distribution of a local (left) and the global (middle and right) minimizers of the discontinuous-forcing optimization problem. Black corresponds to a value of $\rho = 0$ and white corresponds to a value of $\rho = 1$, with the gray regions indicating intermediate values. The arrows indicate the velocity profile of the solutions.

$\rho \in H^1(U_\theta)$ for every $\theta > 0$. In this particular example, the support of ρ is compactly contained in the support of the velocity in all three solutions. Therefore, we conclude that $\rho \in H^1(\Omega)$.

Consider a $(\mathcal{P}_2)^2 \times \mathcal{P}_1$ Taylor–Hood finite element discretization for the velocity and the pressure, and a \mathcal{P}_0 piecewise constant finite element discretization for the material distribution. Since all three solutions are isolated local minimizers, by [Theorem 4](#), there exists a sequence of finite element solutions to the discretized first-order optimality conditions that strongly converges to the figure eight solution, and different sequences of different finite element solutions that strongly converge to the two annulus solutions. Their existence is confirmed in [Fig. 2](#).

Since $\rho \in H^1(\Omega)$ and we are using a \mathcal{P}_0 finite element discretization, a naïve prediction for the convergence rate of the L^2 -norm error of the material distribution is $\mathcal{O}(h)$. This rate is observed in the bottom left panel of [Fig. 2](#). Moreover, since the minimum regularity of the velocity is $\mathbf{u} \in H^2(\Omega)^d$, and we are using a $(\mathcal{P}_2)^2$ finite element discretization, a prediction for the expected convergence rates of the velocity are $\mathcal{O}(h)$ and $\mathcal{O}(h^2)$ for the H^1 -norm and L^2 -norm errors of the velocity, respectively.

In the standard Stokes system, the regularity of \mathbf{u} is related to the regularity of the forcing term $\mathbf{f} \in H^s(\Omega)^d$, such that $\mathbf{u} \in H^{s+2}(\Omega)^d$ (assuming the domain and boundary data are also suitably regular). Here, the regularity of the forcing term satisfies $s < 1/2$. If we assume that the velocity has the additional regularity $\mathbf{u} \in H^{s+2}(\Omega)^d$, $s \in (0, 1/2)$, in this context, a prediction for the upper limit of the convergence rate is $\mathcal{O}(h^r)$ and $\mathcal{O}(h^{t+1})$, for some $r, t \in [1, s+1]$, for the H^1 -norm and L^2 -norm errors of the velocity, respectively. The rates observed in the top panels of [Fig. 2](#) match this prediction. The H^1 -norm error is decreasing at a rate slightly faster than $\mathcal{O}(h^{3/2})$ for all three solutions and the L^2 -norm error convergence rate is $\mathcal{O}(h^2)$ for the figure eight solution and $\mathcal{O}(h^{5/2})$ for the annuli solutions. We hypothesize that the upper limit of the convergence rate of the L^2 -norm error of the velocity is bounded by the relatively slower rate of the convergence of the material distribution.

Finally, since the minimum regularity of the pressure is $p \in H^1(\Omega)$ and the discretization is \mathcal{P}_1 , a prediction for the convergence rate of the L^2 -norm error is $\mathcal{O}(h^r)$, for some $r \in [1, s+1]$. Initially, the convergence rate is $\mathcal{O}(h^{3/2})$ which matches our naïve prediction. However, on finer meshes, the convergence rate increases. We hypothesize that this speedup is artificial and is caused by the lack of resolution of the refined finite element solutions that are being used as proxies for the solutions of the infinite-dimensional problem in the error norm estimate. Qualitatively, it can be checked that mesh refinement in areas where the discretized material distribution lies between $1/10$ and $9/10$ is an ineffective strategy for improving the approximation of the pressure that solves the infinite-dimensional problem over the whole domain.

5.2. Double-pipe

Consider the optimization problem [\(BP\)](#) [6, Sec. 4.5], with two prescribed flow inputs and two prescribed outputs, where $\Omega = (0, 3/2) \times (0, 1)$, $\gamma = 1/3$, $\mathbf{f} = (0, 0)^\top$ and $\nu = 1$ and the boundary conditions on \mathbf{u} are given by the boundary data

$$\mathbf{g}(x, y) = \begin{cases} (1 - 144(y - 3/4)^2, 0)^\top & \text{if } 2/3 \leq y \leq 5/6, x = 0 \text{ or } 3/2, \\ (1 - 144(y - 1/4)^2, 0)^\top & \text{if } 1/6 \leq y \leq 1/3, x = 0 \text{ or } 3/2, \\ (0, 0)^\top & \text{elsewhere on } \partial\Omega. \end{cases} \quad (5.2)$$

The function α is as given in [\(2.1\)](#), with $\bar{\alpha} = 2.5 \times 10^4$ and $q = 1/10$. [Fig. 3](#) visualizes the setup of the problem and depicts the material distribution of the two minimizers. One local minimizer is a straight channel solution and the global minimizer is in the form of a double-ended wrench.

We employ a $(\mathcal{P}_2)^2 \times \mathcal{P}_1$ Taylor–Hood finite element discretization for the velocity and the pressure, and a \mathcal{P}_1 continuous piecewise linear finite element discretization for the material distribution. This example satisfies all the conditions of

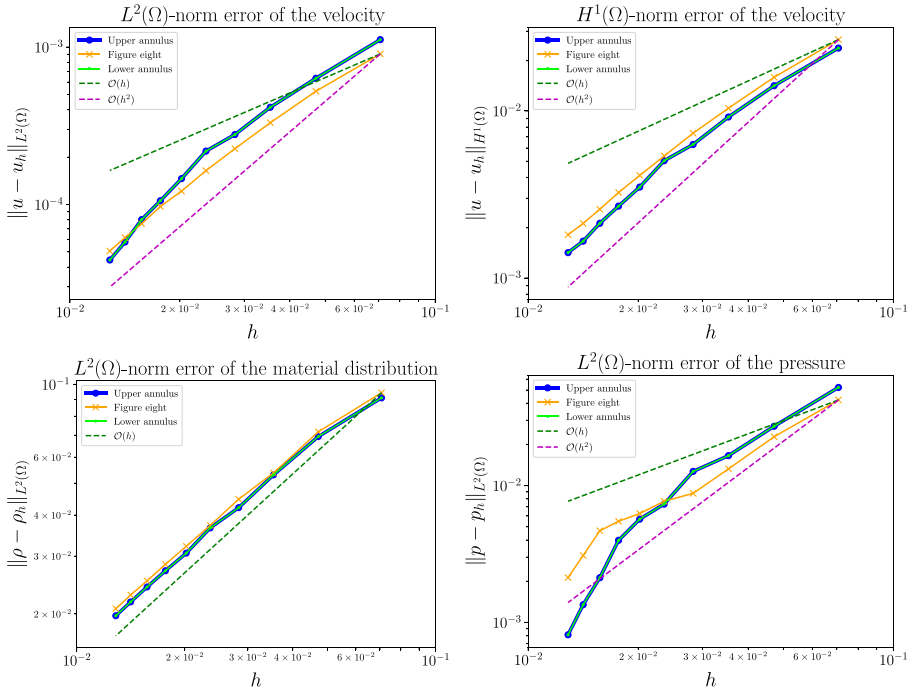


Fig. 2. The convergence of \mathbf{u}_h , ρ_h , and p_h in the discontinuous-forcing problem for the figure eight and annulus solutions on structured meshes, with a $\mathcal{P}_0 \times (\mathcal{P}_2)^2 \times \mathcal{P}_1$ discretization for $(\rho_h, \mathbf{u}_h, p_h)$.



Fig. 3. The setup of the double-pipe problem is shown on the left. The middle and right figures depict the material distribution of the straight channel and double-ended wrench solutions, respectively. Black corresponds to a value of $\rho = 0$ and white corresponds to a value of $\rho = 1$, with the gray regions indicating intermediate values.

Theorem 4 and, for both minimizers, we numerically verify that there exists a sequence of finite element solutions that strongly converges to it in [Fig. 4](#).

As mentioned earlier in this section, it may be possible to find a sequence of mesh sizes, (h_i) , such that there exist two different sequences of finite element solutions that strongly converge to the same isolated minimizer. In [Fig. 5](#), we depict two different straight channel finite element solutions that exist on the same unstructured mesh where $h = 0.04$. Both solutions satisfy the discretized first-order optimality conditions $(\text{FOC1}_h) - (\text{FOC3}_h)$ and both locally minimize $J(\mathbf{v}_h, \eta_h)$. Choosing one over the other would change the convergence pattern of the strongly converging sequence. This may cause difficulty in practice, as optimization strategies are unlikely to discover the discretized global minimum without additional selection mechanisms.

6. Conclusions

In this work we have studied the fluid topology optimization model of Borrvall and Petersson [\[6\]](#). In the case of a homogeneous Dirichlet boundary condition and under a mild convexity assumption on the inverse permeability term, α , we have shown that the material distribution necessarily lives in the Sobolev space H^1 inside any compact subset of the support of the velocity. Moreover, we have formally treated the nonconvexity of the optimization problem (including the case of inhomogeneous Dirichlet boundary conditions) and have shown that, given an isolated minimizer of the infinite-dimensional problem, there exists a sequence of discretized solutions, satisfying the associated first-order optimality

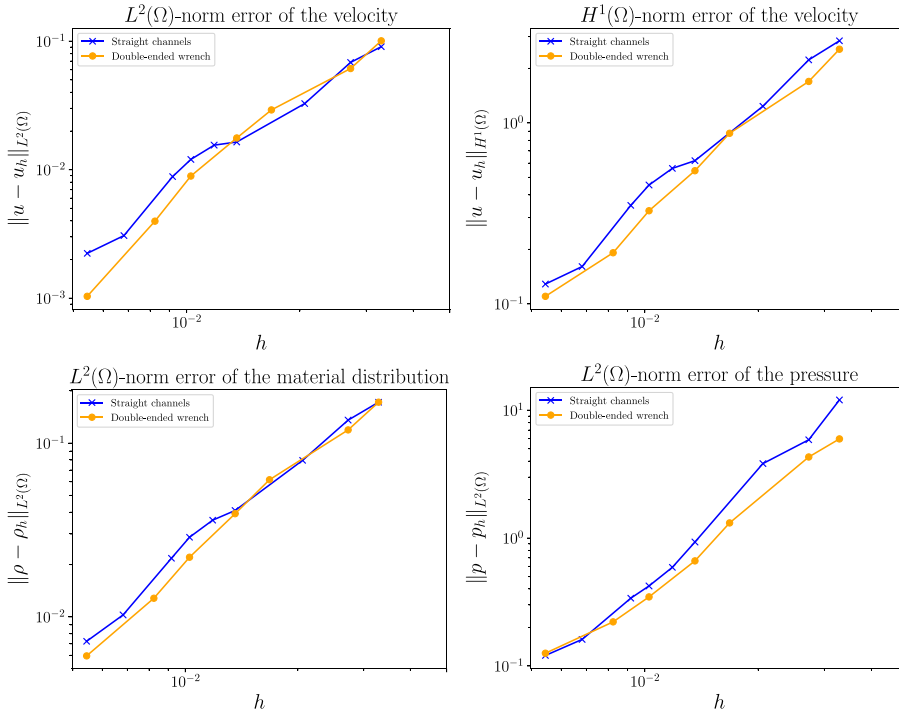


Fig. 4. The convergence of \mathbf{u}_h , ρ_h , and p_h for the double-pipe problem for both the straight channel and double-ended wrench solutions on an unstructured mesh with a $\mathcal{P}_1 \times (\mathcal{P}_2)^2 \times \mathcal{P}_1$ discretization for $(\rho_h, \mathbf{u}_h, p_h)$.

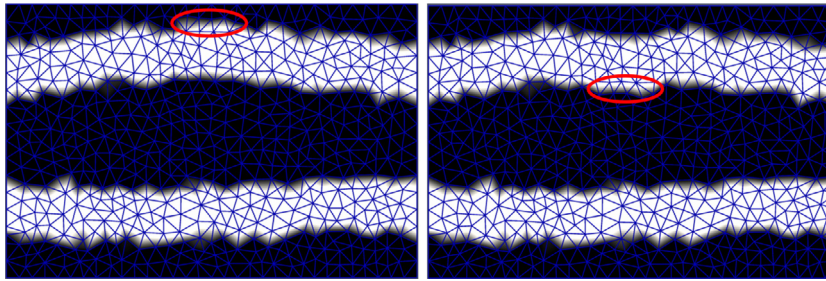


Fig. 5. Two different straight channel finite element solutions of the double-pipe optimization problem that exist on the same unstructured mesh where $h = 0.04$. The differences can be spotted at the midway point of the top channel.

conditions, that *strongly* converges to the minimizer in the appropriate spaces. We have numerically verified that these sequences exist and have discussed the observed convergence rates.

Appendix. Elliptic regularity of the generalized Stokes system with Dirichlet boundary data

Lemma 5. Let the domain Ω be either a convex polygon in two dimensions or a convex polyhedron in three dimensions and consider the triple $(\mathbf{u}, \rho, p) \in H_{\mathbf{g}}^1(\Omega)^d \times C_\gamma \times L_0^2(\Omega)$ that satisfies (FOC1)–(FOC3). Suppose that the forcing term $\mathbf{f} \in L^2(\Omega)^d$ and the boundary datum \mathbf{g} is the boundary trace of a function $\hat{\mathbf{g}} \in H^2(\Omega)^d$ on the boundary $\partial\Omega$ and satisfies $\int_{\partial\Omega} \mathbf{g} \cdot \mathbf{n} \, ds = 0$. Then $\mathbf{u} \in H^2(\Omega)^d$ and $p \in H^1(\Omega)$.

Proof. The idea of the proof is to reduce the system (FOC1)–(FOC2) to the standard Stokes system with a homogeneous Dirichlet boundary condition and invoke the regularity results of Kellogg and Osborn [36] and in three dimensions the results found in Kozlov et al. [37] and Maz'ya and Shaposhnikova [38].

Let $\mathbf{w} := \mathbf{u} - \hat{\mathbf{g}}$. Since the trace operator is a linear operator, we see that $\mathbf{w}|_{\partial\Omega} = (\mathbf{u} - \hat{\mathbf{g}})|_{\partial\Omega} = \mathbf{g} - \mathbf{g} = \mathbf{0}$. Since $\mathbf{u} \in H^1(\Omega)^d$ and $\hat{\mathbf{g}} \in H^2(\Omega)^d$, then $\mathbf{w} \in H_0^1(\Omega)^d$.

By substituting \mathbf{w} into (FOC1)–(FOC2), we see that (FOC1)–(FOC2) is equivalent to finding $(\mathbf{w}, p) \in H_0^1(\Omega)^d \times L_0^2(\Omega)$ that satisfies for all $(\mathbf{v}, q) \in H_0^1(\Omega)^d \times L_0^2(\Omega)$:

$$\int_{\Omega} \nabla \mathbf{w} : \nabla \mathbf{v} - p \operatorname{div}(\mathbf{v}) \, dx = \int_{\Omega} (\mathbf{f} - \alpha(\rho)(\mathbf{w} + \hat{\mathbf{g}})) \cdot \mathbf{v} - \nabla \hat{\mathbf{g}} : \nabla \mathbf{v} \, dx, \quad (\text{A.1})$$

$$\int_{\Omega} q \operatorname{div}(\mathbf{w} + \hat{\mathbf{g}}) \, dx = 0. \quad (\text{A.2})$$

Define $\hat{\mathbf{f}}$ as $\hat{\mathbf{f}} := \mathbf{f} - \alpha(\rho)(\mathbf{w} + \hat{\mathbf{g}}) + \Delta \hat{\mathbf{g}}$. Since $\mathbf{f} \in L^2(\Omega)^d$, $\alpha(\rho) \in L^\infty(\Omega)$, $\mathbf{w} \in H_0^1(\Omega)^d$, and $\hat{\mathbf{g}} \in H^2(\Omega)^d$, then $\hat{\mathbf{f}} \in L^2(\Omega)^d$. By an application of integration by parts on the final term on the right-hand side of (A.1), we see that (A.1)–(A.2) is equivalent to finding $(\mathbf{w}, p) \in H_0^1(\Omega)^d \times L_0^2(\Omega)$ that satisfies for all $(\mathbf{v}, q) \in H_0^1(\Omega)^d \times L_0^2(\Omega)$:

$$\int_{\Omega} \nabla \mathbf{w} : \nabla \mathbf{v} - p \operatorname{div}(\mathbf{v}) \, dx = \int_{\Omega} \hat{\mathbf{f}} \cdot \mathbf{v} \, dx, \quad (\text{A.3})$$

$$\int_{\Omega} q \operatorname{div}(\mathbf{w}) \, dx = \int_{\Omega} q \phi \, dx, \quad (\text{A.4})$$

where $\phi = -\operatorname{div}(\hat{\mathbf{g}})$ a.e. and the divergence theorem implies that

$$\int_{\Omega} \phi \, dx = - \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{n} \, ds = 0. \quad (\text{A.5})$$

We note that (A.3)–(A.4) is the standard Stokes system with a homogeneous Dirichlet boundary condition and forcing term $\hat{\mathbf{f}} \in L^2(\Omega)^d$. Therefore, by the elliptic regularity of the Stokes system [36,37] and [38, Th. 13], $\mathbf{w} \in H^2(\Omega)^d$ and $p \in H^1(\Omega)$. Since $\mathbf{u} = \mathbf{w} + \hat{\mathbf{g}}$ and $\hat{\mathbf{g}} \in H^2(\Omega)^d$, we conclude that $\mathbf{u} \in H^2(\Omega)^d$. \square

References

- [1] L. Adam, M. Hintermüller, D. Peschka, T.M. Surowiec, Optimization of a multiphysics problem in semiconductor laser design, *SIAM J. Appl. Math.* 79 (1) (2019) 257–283, <http://dx.doi.org/10.1137/18M1179183>.
- [2] I.G. Jang, I.Y. Kim, Computational study of Wolff's law with trabecular architecture in the human proximal femur using topology optimization, *J. Biomech.* 41 (11) (2008) 2353–2361, <http://dx.doi.org/10.1016/j.jbiomech.2008.05.037>.
- [3] J. Liu, A.T. Gaynor, S. Chen, Z. Kang, K. Suresh, A. Takezawa, L. Li, J. Kato, J. Tang, C.C.L. Wang, L. Cheng, X. Liang, A.C. To, Current and future trends in topology optimization for additive manufacturing, *Struct. Multidiscip. Optim.* 57 (6) (2018) 2457–2483, <http://dx.doi.org/10.1007/s00158-018-1994-3>.
- [4] G. Allaire, *Shape Optimization by the Homogenization Method*, Vol. 146, Springer Science & Business Media, 2012, <http://dx.doi.org/10.1007/978-1-4684-9286-6>.
- [5] M.P. Bendsøe, O. Sigmund, *Topology Optimization*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, <http://dx.doi.org/10.1007/978-3-662-05086-6>.
- [6] T. Borrvall, J. Petersson, Topology optimization of fluids in Stokes flow, *Internat. J. Numer. Methods Fluids* 41 (1) (2003) 77–107, <http://dx.doi.org/10.1002/flid.426>.
- [7] J. Alexandersen, C.S. Andreasen, A review of topology optimisation for fluid-based problems, *Fluids* 5 (1) (2020) 29, <http://dx.doi.org/10.3390/fluids5010029>.
- [8] D.H. Alonso, L.F.N. de Sá, J.S.R. Saenz, E.C.N. Silva, Topology optimization applied to the design of 2D swirl flow devices, *Struct. Multidiscip. Optim.* 58 (6) (2018) 2341–2364, <http://dx.doi.org/10.1007/s00158-018-2078-0>.
- [9] D.H. Alonso, J.S.R. Saenz, E.C.N. Silva, Non-newtonian laminar 2D swirl flow design by the topology optimization method, *Struct. Multidiscip. Optim.* (2020) 1–23, <http://dx.doi.org/10.1007/s00158-020-02499-2>.
- [10] L.H. Olesen, F. Okkels, H. Bruus, A high-level programming-language implementation of topology optimization applied to steady-state Navier–Stokes flow, *Internat. J. Numer. Methods Engrg.* 65 (7) (2006) 975–1001, <http://dx.doi.org/10.1002/nme.1468>.
- [11] A. Gersborg-Hansen, O. Sigmund, R.B. Haber, Topology optimization of channel flow problems, *Struct. Multidiscip. Optim.* 30 (3) (2005) 181–192, <http://dx.doi.org/10.1007/s00158-004-0508-7>.
- [12] A. Evgrafov, State space Newton's method for topology optimization, *Comput. Methods Appl. Mech. Engrg.* 278 (2014) 272–290, <http://dx.doi.org/10.1016/j.cma.2014.06.005>.
- [13] A. Evgrafov, Topology optimization of slightly compressible fluids, *ZAMM J. Appl. Math. Mech./Z. Angew. Math. Mech.: Appl. Math. Mech.* 86 (1) (2006) 46–62, <http://dx.doi.org/10.1002/zamm.200410223>.
- [14] Y. Deng, Y. Wu, Z. Liu, *Topology Optimization Theory for Laminar Flow: Applications in Inverse Design of Microfluidics*, Springer Singapore, Singapore, 2018, <http://dx.doi.org/10.1007/978-981-10-4687-2>.
- [15] J.K. Guest, J.H. Prévost, Topology optimization of creeping fluid flows using a Darcy–Stokes finite element, *Internat. J. Numer. Methods Engrg.* 66 (3) (2006) 461–484, <http://dx.doi.org/10.1002/nme.1560>.
- [16] N. Aage, T.H. Poulsen, A. Gersborg-Hansen, O. Sigmund, Topology optimization of large scale Stokes flow problems, *Struct. Multidiscip. Optim.* 35 (2) (2008) 175–180, <http://dx.doi.org/10.1007/s00158-007-0128-0>.
- [17] S. Kreissl, G. Pingen, K. Maute, Topology optimization for unsteady flow, *Internat. J. Numer. Methods Engrg.* 87 (13) (2011) 1229–1253, <http://dx.doi.org/10.1002/nme.3151>.
- [18] L.F.N. Sá, J.S. Romero, O. Horikawa, E.C.N. Silva, Topology optimization applied to the development of small scale pump, *Struct. Multidiscip. Optim.* 57 (5) (2018) 2045–2059, <http://dx.doi.org/10.1007/s00158-018-1966-7>.
- [19] H. Garcke, C. Hecht, A phase field approach for shape and topology optimization in Stokes flow, in: *New Trends in Shape Optimization*, Springer, 2015, pp. 103–115, http://dx.doi.org/10.1007/978-3-319-17563-8_5.
- [20] H. Garcke, C. Hecht, Applying a phase field approach for shape optimization of a stationary Navier–Stokes flow, *ESAIM Control Optim. Calc. Var.* 22 (2) (2016) 309–337, <http://dx.doi.org/10.1051/cocv/2015006>.
- [21] H. Garcke, C. Hecht, M. Hinze, C. Kahle, Numerical approximation of phase field based shape and topology optimization for fluids, *SIAM J. Sci. Comput.* 37 (4) (2015) A1846–A1871, <http://dx.doi.org/10.1137/140969269>.

- [22] C.-J. Thore, Topology optimization of Stokes flow with traction boundary conditions using low-order finite elements, *Comput. Methods Appl. Mech. Engrg.* 386 (2021) 114177, <http://dx.doi.org/10.1016/j.cma.2021.114177>.
- [23] L.C. Evans, *Partial Differential Equations*, second ed., American Mathematical Society, 2010.
- [24] I.P.A. Papadopoulos, Computing Multiple Solutions of Topology Optimization Problems (Ph.D. thesis), University of Oxford, 2021, URL <https://ora.ox.ac.uk/objects/uuid:455f5465-af68-49c7-8e11-96909cdb82a1>.
- [25] W.S. Ożański, The Lagrange multiplier and the stationary Stokes equations, *J. Appl. Anal.* 23 (2) (2017) 137–140, <http://dx.doi.org/10.1515/jaa-2017-0017>.
- [26] V. Girault, P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Vol. 5, Springer-Verlag Berlin Heidelberg, 1986, <http://dx.doi.org/10.1007/978-3-642-61623-5>.
- [27] J. Petersson, A finite element analysis of optimal variable thickness sheets, *SIAM J. Numer. Anal.* 36 (6) (1999) 1759–1778, <http://dx.doi.org/10.1137/S0036142996313968>.
- [28] A. Logg, K.-A. Mardal, G. Wells, *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*, Vol. 84, Springer Science & Business Media, 2012.
- [29] I.P.A. Papadopoulos, P.E. Farrell, T.M. Surowiec, Computing multiple solutions of topology optimization problems, *SIAM J. Sci. Comput.* 43 (3) (2021) A1555–A1582, <http://dx.doi.org/10.1137/20M1326209>.
- [30] S.J. Benson, T.S. Munson, Flexible complementarity solvers for large-scale applications, *Optim. Methods Softw.* 21 (1) (2003) 155–168, <http://dx.doi.org/10.1080/10556780500065382>.
- [31] P.E. Farrell, A. Birkisson, S.W. Funke, Deflation techniques for finding distinct solutions of nonlinear partial differential equations, *SIAM J. Sci. Comput.* 37 (4) (2015) A2026–A2045, <http://dx.doi.org/10.1137/140984798>.
- [32] P.E. Farrell, M. Croci, T.M. Surowiec, Deflation for semismooth equations, *Optim. Methods Softw.* (2019) 1–24, <http://dx.doi.org/10.1080/10556788.2019.1613655>.
- [33] P.R. Amestoy, I.S. Duff, J.-Y. L'Excellent, J. Koster, A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM J. Matrix Anal. Appl.* (2001) <http://dx.doi.org/10.1137/S0895479899358194>.
- [34] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. Gropp, R. Tran Mills, T. Munson, K. Rupp, P. Sana, B. Smith, S. Zampini, H. Zhang, H. Zhang, PETSc Users Manual, Tech. Rep. ANL-95/11 - Revision 3.11, Argonne National Laboratory, 2019, URL <http://www.mcs.anl.gov/petsc>.
- [35] Software used in 'Numerical analysis of a topology optimization problem for Stokes flow', 2021, <http://dx.doi.org/10.5281/zenodo.4514054>.
- [36] R.B. Kellogg, J.E. Osborn, A regularity result for the Stokes problem in a convex polygon, *J. Funct. Anal.* 21 (4) (1976) 397–431, [http://dx.doi.org/10.1016/0022-1236\(76\)90035-5](http://dx.doi.org/10.1016/0022-1236(76)90035-5).
- [37] V.A. Kozlov, V.G. Maz'ya, C. Schwab, On singularities of solutions to the Dirichlet problem of hydrodynamics near the vertex of a cone, *J. R. Angew. Math.* 456 (1994) 65–97, <http://dx.doi.org/10.1515/crll.1994.456.65>.
- [38] V. Maz'ya, T. Shaposhnikova, Recent progress in elliptic equations and systems of arbitrary order with rough coefficients in Lipschitz domains, *Bull. Math. Sci.* 1 (1) (2011) 33–77, <http://dx.doi.org/10.1007/s13373-011-0003-6>.