COMPUTING MULTIPLE SOLUTIONS OF TOPOLOGY OPTIMIZATION PROBLEMS *

IOANNIS P. A. PAPADOPOULOS†, PATRICK E. FARRELL‡, AND THOMAS M. SUROWIEC §

Abstract. Topology optimization problems often support multiple local minima due to a lack of convexity. Typically, gradient-based techniques combined with continuation in model parameters are used to promote convergence to more optimal solutions; however, these methods can fail even in the simplest cases. In this paper, we present an algorithm to perform a systematic exploratory search for the solutions of the optimization problem via second-order methods without a good initial guess. The algorithm combines the techniques of deflation, barrier methods and primal-dual active set solvers in a novel way. We demonstrate this approach on several numerical examples, observe mesh-independence in certain cases and show that multiple distinct local minima can be recovered.

Key words. topology optimization, deflation, barrier methods, second-order methods

AMS subject classifications. 35Q35, 49M15, 65K05, 65K10, 74P05, 74P10, 90C26, 90C51

1. Introduction.

Topology optimization has become popular as an effective technique in structural and additive manufacturing, and has found uses in architecture, medicine and material science [2, 31, 35]. The objective is to find the optimal distribution of a fluid or material within a given domain that minimizes a problem-specific cost functional. In contrast to shape optimization, the topology of the structure does not need to be chosen a priori.

There are several mathematical parametrizations for the topology of a material including density approaches [7, 8, 11, 39] and level set methods [3, 4, 59]; these can be optimized by a variety of strategies such as topological derivatives [53], evolutionary methods [63], the method of moving asymptotes [55], and barrier methods [20, 30, 37, 47]. We choose to represent our topology with the density approach. This introduces a function, denoted ρ, that represents the material distribution over the given domain. Ideally we would find an optimizing material distribution $\rho : \Omega \to \{0, 1\}$ indicating presence or absence of material. However, this is numerically intractable in general and we therefore consider densities $\rho : \Omega \to [0, 1]$ in order to exploit continuous optimization techniques. The model is then regularized to favor solutions where ρ is close to zero or one.

Due to the nonlinear relation between ρ and the solution of the underlying physical system, multiple local minima can occur even in problems with a linear governing partial differential equation (PDE). For example, minimizing the power dissipation of a fluid governed by the Stokes equations flowing through a pipe can give rise to distinct pipe configurations that locally minimize the power lost to dissipation [11, Sec. 4.5]. Currently, the main technique to address this is the use of *continuation methods* to

*Submitted to the editors, March 18 2020.

Funding: The first author is supported by the EPSRC Centre for Doctoral Training in Partial Differential Equations: Analysis and Applications [grant number EP/L015811/1] and The MathWorks, Inc. The second author is supported by the Engineering and Physical Sciences Research Council [grant number EP/R029423/1]. The third author is supported by the German Research Foundation [DFG-Grant SU 963/1-1]

†Mathematical Institute, University of Oxford, Oxford, UK (ioannis.papadopoulos@maths.ox.ac.uk.

‡Mathematical Institute, University of Oxford, Oxford, UK (patrick.farrell@maths.ox.ac.uk).

§Department of Mathematics and Computer Science, Philipps-Universität Marburg, Marburg, DE (surowiec@mathematik.uni-marburg.de).

promote convergence to better local minima. However, Stolpe and Svanberg [54] have provided elementary examples where these continuation methods fail. For example, a solid isotropic material with penalization (SIMP) formulation [8] of the compliance minimization of a six-bar truss can be reduced to the optimization problem [54, Sec. 3.1],

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \left(\max\{ \frac{8\beta_t}{x_1^{p_s}+5x_2^{p_s}} + \frac{2\beta_t}{5x_1^{p_s}+x_2^{p_s}}, \frac{8}{x_1^{p_s}+5x_2^{p_s}} + \frac{18}{5x_1^{p_s}+x_2^{p_s}} \} \right)$$
$$\text{such that } \; x_1 + x_2 = 1, \quad 0 \le x_1, x_2 \le 1.$$

Here p_s denotes the SIMP continuation parameter and $\beta_t = 2(1-\nu_t^2)/E$, where ν_t is the Poisson ratio and E is the modulus of elasticity. SIMP is used to penalize solutions that are not either zero or one and is further discussed in subsection 2.3. A typical strategy is to find a minimizer to the optimization problem at $p_s = 1$, and then at each continuation step use the previous solution as initial guess for the next value of p_s. In this case, suppose we fix $\beta_t = 2.6$. A poor starting guess for $p_s = 1$ can converge to the local minimum $x = (0.5, 0.5)$. Then even as $p_s \to \infty$, the continuation method will always return $x = (0.5, 0.5)$ and will not converge to the true global solution, $x = (0, 1)$.

The calculation of multiple stationary points is important because iterative methods often give no guarantee whether they converge to a local or global minimum. By finding multiple stationary points, one is able to choose the best available, in a post-processing step. Furthermore, an iterative method may converge to a stationary point which is undesirable due to manufacturing or aesthetic reasons; thus industrial applications can benefit from having a choice of multiple locally optimal configurations [18].

In this paper we formulate an algorithm, which we call the *deflated barrier method*, for finding multiple stationary points of topology optimization problems and present several large-scale numerical examples arising from the finite element discretization of PDEs. An example we consider is the topology optimization of the power dissipation of fluid flow governed by the incompressible Navier–Stokes equations on a rectangular domain with five small decagonal holes. We discover 42 stationary points of this optimization problem with the deflated barrier method. The material distribution of these solutions are shown in Figure 1.

The deflated barrier method is a combination of *deflation* [13, 21, 22], *barrier methods* [23, 24, 25, 48, 49, 58, 60], *primal-dual active set solvers* [9, 28] and *predictor-corrector methods* [51]. The combination of primal-dual active set solvers, barrier and deflation methods in the manner proposed is novel. The combination does not suffer the poor behavior that barrier methods traditionally exhibit as the barrier parameter approaches zero. In fact, in our numerical examples, the combination performs better than the optimize-then-discretize formulation of the primal-dual interior point method where Newton–Kantorovich iterates are used to solve the subproblems, either approximately or exactly. The predictor-corrector method is also adapted for use with box-constrained variables to ensure the predictor is feasible. The main contribution of this work is an algorithm to robustly determine multiple solutions to nonconvex, inequality and box-constrained infinite-dimensional optimization problems starting from poor initial guesses.

Other approaches to computing multiple solutions of topology optimization problems are possible. Zhang and Norato [64] apply the *tunneling* method [33] to these problems, adapting the method of moving asymptotes. Tunneling proceeds by finding

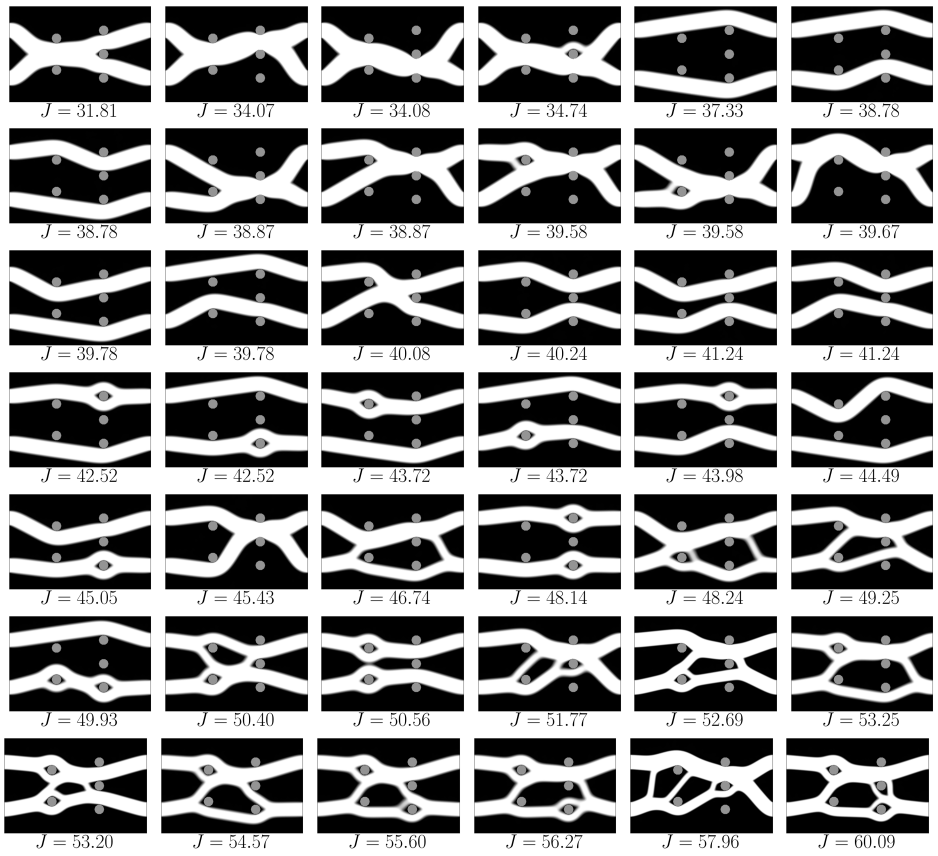


FIG. 1. *The material distribution of 42 stationary points of the five-holes double-pipe optimization problem as discovered by the deflated barrier method, and their associated energies J. The fluid flow is governed by the incompressible Navier–Stokes equations. The formulation of the problem is described in subsection 4.4. Black corresponds to a value of $\rho = 0$, white corresponds to a value of $\rho = 1$, and the gray regions are the five small holes.*

a single minimum, then looking for other controls that yield the same functional value (attempting to tunnel into other basins) by solving an auxiliary equation. Deflation is used in the tunneling phase to ensure that the Gauss–Newton procedure applied to the tunneling functional does not converge to the current state.

The outline of the paper is as follows. In section 2 we formulate some topology optimization problems for pipe design and structural compliance. The deflated barrier method is described in section 3. Several examples of topology optimization problems are given in section 4, where we discover multiple solutions for Navier–Stokes flow, Stokes flow, and structural compliance, and consider the performance of our algorithm. In section 5 we outline our conclusions. A result concerning the equivalence of Hintermüller et al.'s primal-dual active set strategy [28] and Benson and Munson's reduced space active-set strategy [9] is given in Appendix A. In Appendix B we describe our novel feasible tangent prediction method.

2. Topology optimization formulations.

2.1. Topology optimization of Stokes flow. We consider the formulation of the topology optimization of fluids proposed in the pioneering work of Borrvall and Petersson [11]. They derive a 'generalized Stokes problem' incorporating a material distribution variable which has a value of one where fluid is present and zero where there is void. The derived optimization problem requires no further regularization for well-posedness, in contrast to structural topology optimization. The optimization problem supports (not necessarily unique) local minima.

The topology optimization problem of Borrvall and Petersson is

$$\text{(BP)} \qquad \min_{(\boldsymbol{u},\rho)\in H^1_{\boldsymbol{g},\mathrm{div}}(\Omega)^d\times C_\gamma} J(\boldsymbol{u},\rho) := \frac{1}{2}\int_\Omega \left(\alpha(\rho)|\boldsymbol{u}|^2 + \nu|\nabla\boldsymbol{u}|^2 - 2\boldsymbol{f}\cdot\boldsymbol{u}\right)\mathrm{d}x,$$

where $\boldsymbol{u}$ denotes the velocity of the fluid, ρ is the material distribution of the fluid and

$$H^1_{\boldsymbol{g}}(\Omega)^d := \{\boldsymbol{v} \in \underbrace{H^1(\Omega)\times\cdots\times H^1(\Omega)}_{d \text{ times}} : \boldsymbol{v}|_{\partial\Omega} = \boldsymbol{g}\},$$

$$H^1_{\boldsymbol{g},\mathrm{div}}(\Omega)^d := \{\boldsymbol{v} \in H^1_{\boldsymbol{g}}(\Omega)^d : \mathrm{div}(\boldsymbol{v}) = 0 \text{ a.e. in } \Omega\},$$

$$C_\gamma := \left\{\eta \in L^\infty(\Omega) : 0 \le \eta \le 1 \text{ a.e.}, \ \int_\Omega \eta\,\mathrm{d}x \le \gamma|\Omega|, \ \gamma \in (0,1)\right\}.$$

In this work, $H^1(\Omega)$ denotes the Sobolev space $W^{1,2}(\Omega)$ and $L^\infty(\Omega)$ denotes the vector space of essentially bounded measurable functions equipped with the essential supremum norm. Furthermore, $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain with dimension $d = 2$ or $d = 3$, $\boldsymbol{f} \in L^2(\Omega)^d$ is a body force and $\nu > 0$ is the (constant) viscosity. The restriction, $|_{\partial\Omega}$, is to be understood in the boundary trace sense. Moreover, the boundary data $\boldsymbol{g} \in H^{1/2}(\partial\Omega)^d$ and $\boldsymbol{g} = \boldsymbol{0}$ on $\Gamma \subset \partial\Omega$, with $\mathcal{H}^{d-1}(\Gamma) > 0$, i.e. Γ has nonzero Hausdorff measure on the boundary. Mixed boundary conditions are discussed in subsection 4.2. 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 satisfies the following properties:

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

generating a superposition operator also denoted $\alpha : C_\gamma \to L^\infty(\Omega; [\underline{\alpha}, \overline{\alpha}])$. Typically, in the literature α takes the form [11, 20]

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

where $q > 0$ is a penalty parameter, so that $\lim_{q\to\infty}\alpha(\rho) = \bar{\alpha}(1-\rho)$. The objective functional (BP) can be interpreted as the total potential power of the flow. The first and second terms in the integral measure the energy lost by the flow through the porous medium and the energy lost due to viscous dissipation, respectively. The third term attempts to maximize the flow velocities at the applied body force. (BP) is discussed in further detail by Borrvall and Petersson [11].

Remark 2.1. 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 $\alpha(\rho) : \Omega \to [\underline{\alpha}, \overline{\alpha}]$ is Lebesgue measurable.

THEOREM 2.2. *[11, Th. 3.1] Suppose that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain, with $d = 2$ or $d = 3$ and α satisfies properties (A1)–(A3). Then there exists a pair $(\boldsymbol{u}, \rho) \in H^1_{\boldsymbol{g},\mathrm{div}}(\Omega) \times C_\gamma$ that minimizes J, as defined in (BP).*

Due to the lack of strict convexity in (BP), a minimizing pair is not necessarily unique.

2.2. Construction of the barrier functional. In this subsection we formulate a barrier functional with an enlarged feasible set that will be employed by our algorithm to find multiple solutions of the Borrvall–Petersson optimization problem.

We first consider the volume constraint. This constraint is typically modeled as an inequality constraint. However, as we show below, this constraint is active at an optimal solution, and so we may also apply it as an equality constraint. To the best of our knowledge, the following result is novel.

PROPOSITION 2.3. *If the pair $(\boldsymbol{u}_*, \rho_*)$ is an isolated local or global minimizer of J as defined in (BP) and $\gamma < 1$, then $\int_\Omega \rho_* \, \mathrm{d}x = \gamma|\Omega|$.*

Proof by contradiction. Suppose there exists a pair $(\boldsymbol{u}_*, \rho_*) \in H^1_{\boldsymbol{g},\mathrm{div}}(\Omega)^d \times C_\gamma$ that is an isolated local or global minimizer of $J(\boldsymbol{u}, \rho)$ such that $V := \int_\Omega \rho_* \, \mathrm{d}x < \gamma|\Omega|$. By the definition of an isolated local minimizer, there exists an $r > 0$ such that for any $(\boldsymbol{v}, \eta)$ that satisfies,

$$\|\boldsymbol{u}_* - \boldsymbol{v}\|_{H^1(\Omega)} + \|\rho_* - \eta\|_{L^\infty(\Omega)} \leq r$$

then $J(\boldsymbol{u}_*, \rho_*) < J(\boldsymbol{v}, \eta)$. Then for any function $\delta\rho \in C_\gamma$ such that

$$0 < \|\delta\rho\|_{L^1(\Omega)} \leq (\gamma|\Omega| - V), \tag{2.2}$$

$$0 < \|\delta\rho\|_{L^\infty(\Omega)} \leq r, \tag{2.3}$$

$$0 \leq \rho_* + \delta\rho \leq 1, \tag{2.4}$$

we have that $\rho_* + \delta\rho \in C_\gamma$ from (2.2) and (2.4) and $\rho_* + \delta\rho$ lies in the L^∞-r-neighborhood of ρ_* from (2.3). Such a $\delta\rho$ exists, for example,

$$\delta\rho = c(1 - \rho_*), \quad \text{where } c = \min\left\{ \frac{r}{\|1 - \rho_*\|_{L^\infty(\Omega)}}, \frac{\gamma|\Omega| - V}{|\Omega| - V} \right\}.$$

We see that $c > 0$ since $r > 0$ and $V < \gamma|\Omega| < |\Omega|$. Furthermore $\delta\rho$ satisfies (2.2)–(2.4) since,

$$\|\delta\rho\|_{L^1(\Omega)} = c \int_\Omega (1 - \rho_*) \mathrm{d}x \leq c(|\Omega| - V) \leq \gamma|\Omega| - V,$$

$$\|\delta\rho\|_{L^\infty(\Omega)} \leq c\|1 - \rho_*\|_{L^\infty(\Omega)} \leq r,$$

$$0 \leq \rho_* + \delta\rho = \rho_* + c(1 - \rho_*) \leq \rho_* + 1 - \rho_* \leq 1.$$

Since $\alpha(\cdot)$ is monotonically decreasing and ρ_* and $\delta\rho$ are non-negative and not equal to zero, then $\alpha(\rho_* + \delta\rho) \leq \alpha(\rho_*)$ a.e. and hence $J(\boldsymbol{u}_*, \rho_* + \delta\rho) \leq J(\boldsymbol{u}_*, \rho_*)$. ∎

Given we can tighten the inequality volume constraint to an equality volume constraint, we now define the Lagrangian and the enlarged feasible-set barrier functional, respectively, as:

$$L(\boldsymbol{u}, \rho, p, p_0, \lambda) := J(\boldsymbol{u}, \rho) - \int_\Omega p \, \mathrm{div}(\boldsymbol{u}) \mathrm{d}x - \int_\Omega p_0 p \, \mathrm{d}x - \int_\Omega \lambda(\gamma - \rho) \mathrm{d}x; \tag{2.5}$$

$$L^{\epsilon_{\log}}_\mu(\boldsymbol{u}, \rho, p, p_0, \lambda) := L(\boldsymbol{u}, \rho, p, p_0, \lambda) - \mu \int_\Omega (\log(\rho + \epsilon_{\log}) + \log(1 + \epsilon_{\log} - \rho)) \mathrm{d}x, \tag{2.6}$$

where $p \in L^2(\Omega)$ denotes the pressure, λ is the Lagrange multiplier for the volume constraint, $p_0 \in \mathbb{R}$ is the Lagrange multiplier to fix the integral of the pressure, $0 \leq \epsilon_{\log} \ll 1$ and $\mu \geq 0$, where μ is the barrier parameter.

The classical barrier functional is given by L^0_μ. The role of $\epsilon_{\log}$ is to enlarge the feasible region permitted by the barrier terms. In the deflated barrier method we do *not* use the barrier terms to enforce the box-constraints on ρ, but rather to perform continuation in the barrier parameter to follow a central path. This provides robust convergence and offers an opportunity to find other solutions of the optimization problem, as explained in section 3.

We note that the Euler–Lagrange equation of $J(\boldsymbol{u}, \rho)$ with respect to $\boldsymbol{u}$ satisfies the generalized Stokes momentum equation formulated by Borrvall and Petersson [11, Eq. 12]. Hence, we are only required to enforce the incompressibility and volume constraints. In the case where we wish to minimize the power dissipation of a fluid flow governed by a generalized Navier–Stokes momentum equation, we are required to introduce three extra Lagrange multipliers, as done in subsection 4.4.

2.3. Topology optimization of the compliance of elastic structures. A significant portion of the topology optimization literature focuses on minimizing the compliance of a structure, such as a Messerschmitt–Bölkow–Blohm (MBB) beam or a cantilever. Compliance problems involve finding the optimal topology of a structure obeying a volume constraint within a specified domain that minimizes the displacement of the structure under a body or boundary force. For simplicity we consider structures that obey linear elasticity. The optimization problem we consider is posed as follows,

$$\text{(C)} \qquad \min_{(\boldsymbol{u},\rho)\in H^1_{\Gamma_D}(\Omega)^d \times C_\gamma} J(\boldsymbol{u}, \rho) := \int_{\Gamma_N} \boldsymbol{f} \cdot \boldsymbol{u} \, \mathrm{d}s$$

such that,

$$\begin{aligned} -\mathrm{div}\,(\sigma) &= 0 && \text{in } \Omega, \\ \sigma &= k(\rho)\left[2\mu_l \varepsilon(\boldsymbol{u}) + \lambda_l \mathrm{tr}(\varepsilon(\boldsymbol{u}))\mathbb{I}\right] && \text{in } \Omega, \\ \sigma\boldsymbol{n} &= \boldsymbol{f} \text{ on } \Gamma_N, \quad 0 \leq \rho \leq 1 \text{ a.e. in } \Omega, \quad \text{and} \quad \int_\Omega \rho \, \mathrm{d}x = \gamma|\Omega|, \end{aligned}$$

where, $H^1_{\Gamma_D}(\Omega)^d := \left\{\boldsymbol{v} \in H^1(\Omega)^d : \boldsymbol{v}|_{\Gamma_D} = \boldsymbol{0}\right\}$, $|_{\Gamma_D}$ is understood in the boundary trace sense, $\boldsymbol{u} = \boldsymbol{u}(\rho)$ denotes the displacement of the structure, σ denotes the stress tensor, the traction $\boldsymbol{f} \in H^{1/2}(\Gamma_N)^d$ is given, $\Gamma_N, \Gamma_D \subset \partial\Omega$ are known boundaries on $\partial\Omega$ such that $\Gamma_N \cup \Gamma_D = \partial\Omega$, μ_l and λ_l are the Lamé coefficients, $\mathrm{tr}(\cdot)$ is the matrix-trace operator, $\mathbb{I}$ is the $d \times d$ identity matrix, $\boldsymbol{n}$ is the outward normal and

$$\varepsilon(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^\top), \quad k(\rho) = \epsilon_{\mathrm{SIMP}} + (1 - \epsilon_{\mathrm{SIMP}})\rho^{p_s},$$

where $0 < \epsilon_{\mathrm{SIMP}} \ll 1$ and $p_s \geq 1$. Unless stated otherwise, we choose $\epsilon_{\mathrm{SIMP}} = 10^{-5}$ and $p_s = 3$. The use of $k(\rho)$ is known as the Solid Isotropic Material with Penalization (SIMP) model. Bendsøe and Sigmund [8, Ch. 1] provide a concise physical interpretation of the SIMP model. In essence, for ρ close to one, $k(\rho)$ is close to one, indicating the presence of material, whereas where ρ is close to zero, $k(\rho)$ approaches ϵ_{SIMP}, indicating void. Thus, k is the reverse of the inverse permeability, α. It is typical to raise ρ to the power of $p_s > 1$ in order to penalize intermediate values of ρ.

We introduce a Lagrange multiplier $\boldsymbol{v} \in H^1_{\Gamma_D}(\Omega)^d$ and reformulate (C) as finding the stationary points $(\boldsymbol{u}, \rho, \boldsymbol{v})$ of

$$\int_{\Gamma_N} \boldsymbol{f} \cdot \boldsymbol{u} \, \mathrm{d}s + \int_\Omega k(\rho) \left[2\mu_l \varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{v}) + \lambda_l \mathrm{tr}(\varepsilon(\boldsymbol{u})) \cdot \mathrm{tr}(\varepsilon(\boldsymbol{v})) \right] \mathrm{d}x - \int_{\Gamma_N} \boldsymbol{f} \cdot \boldsymbol{v} \, \mathrm{d}s \tag{2.7}$$

such that $0 \leq \rho \leq 1$ a.e. in Ω, and $\int_\Omega \rho \, \mathrm{d}x = \gamma |\Omega|$.

By deriving the Euler–Lagrange equations of (2.7), we see that the linear elasticity PDE constraint on $\boldsymbol{u}$ must be satisfied. However, if we consider the adjoint equation involving $\boldsymbol{v}$, it can be verified that $\boldsymbol{v} = -\boldsymbol{u}$. Substituting this relation into (2.7), we see that (2.7) is equivalent to finding the stationary points of

$$2 \int_{\Gamma_N} \boldsymbol{f} \cdot \boldsymbol{u} \, \mathrm{d}s - \int_\Omega k(\rho) \left[2\mu_l \varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{u}) + \lambda_l \mathrm{tr}(\varepsilon(\boldsymbol{u})) \cdot \mathrm{tr}(\varepsilon(\boldsymbol{u})) \right] \mathrm{d}x \tag{2.8}$$

such that $0 \leq \rho \leq 1$ a.e. in Ω, and $\int_\Omega \rho \, \mathrm{d}x = \gamma |\Omega|$. The substitution is useful as it greatly reduces the size of the problem after discretization.

Unfortunately, the problem in general is ill-posed and does not have minimizers in the continuous setting. Naïve attempts at finding minimizers often yield checkerboard patterns of ρ. Although a different choice of finite element spaces may avoid the checkerboarding, the solutions will still be mesh-dependent. As the mesh is refined, the beams of the solutions will become ever thinner, leading to nonphysical solutions in the limit. There are several schemes employed by the topology optimization community to obtain physically reasonable solutions for ρ and they are known as *restriction methods* [8]. We opt for the addition of a Ginzburg–Landau energy term,

$$J_{\mathrm{GL}}(\boldsymbol{u}, \rho) := J(\boldsymbol{u}, \rho) + \frac{\beta\epsilon}{2} \int_\Omega |\nabla \rho|^2 \, \mathrm{d}x + \frac{\beta}{2\epsilon} \int_\Omega \rho(1-\rho) \mathrm{d}x,$$

with $0 < \beta \ll 1$, $0 < \epsilon \ll 1$, to the objective function. J_{GL} requires ρ to be weakly differentiable. Hence we now seek a solution $\rho \in C_\gamma \cap H^1(\Omega)$. Physically, the Ginzburg–Landau term corresponds to penalizing fluctuations in the values of ρ. As $\epsilon \to 0$, it was shown by Modica [40] that the Ginzburg–Landau energy Γ-converges to the perimeter functional associated with restricting $\rho(x) \in \{0, 1\}$, providing rigorous mathematical grounding for this choice of regularization. For sufficiently large values of β, this introduces minima and removes the checkerboarding effect. Other restriction methods used by the topology optimization community include gradient control [10], perimeter constraints [10], sensitivity filtering [12, 52], design filtering [15, 32] and regularized penalty [10].

After these manipulations, the Lagrangian is given by

$$\begin{aligned} L(\boldsymbol{u}, \rho, \lambda) := {} & 2 \int_{\Gamma_N} \boldsymbol{f} \cdot \boldsymbol{u} \, \mathrm{d}s - \int_\Omega k(\rho) \left[2\mu_l \varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{u}) + \lambda_l \mathrm{tr}(\varepsilon(\boldsymbol{u})) \cdot \mathrm{tr}(\varepsilon(\boldsymbol{u})) \right] \mathrm{d}x \\ & + \frac{\beta\epsilon}{2} \int_\Omega |\nabla \rho|^2 \mathrm{d}x + \frac{\beta}{2\epsilon} \int_\Omega \rho(1-\rho) \mathrm{d}x - \int_\Omega \lambda(\gamma - \rho) \mathrm{d}x, \end{aligned}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier for the equality volume constraint. We then define the enlarged feasible-set barrier functional as in (2.6).

We have formulated enlarged feasible-set barrier functionals for both Borrvall–Petersson and structural compliance optimization problems. Finding stationary points of these barrier functionals is equivalent to computing minima, maxima and saddle points of the underlying optimization problems. In the next section we will introduce our algorithm and explain how we obtain multiple stationary points.

3. The deflated barrier method. In the following sections, we describe the components of the deflated barrier method. More specifically, we justify the usage of a barrier method where the subproblems are solved with a primal-dual active set solver to handle the effects of the barrier parameter in the Hessian. This is in contrast to a direct application of a discretize-then-optimize (DTO) primal-dual interior method, which does not use the structure of the original infinite-dimensional optimization problem. In the context of PDE-constrained optimization, ignoring the problem structure often results in mesh-dependence of the solver. Mesh-dependence is the phenomenon whereby with each refinement of the mesh, the number of iterations required by the optimization algorithm increases in an unbounded way [50].

3.1. Choosing a solver for the subproblems. Approximately solving the first order conditions of L^0_μ as $\mu \to 0$ is the classical primal interior point approach to finding the minima of (BP) and (C). Without additional care, a direct implementation results in the following poor numerical behavior:

(B1) The Hessian of $L^0_{\mu_k}(\boldsymbol{z})$ has condition number $\mathcal{O}(1/\mu_k)$. Hence as μ decreases, the computed Newton updates may become inaccurate and require more solver time [24, Th. 4.2];

(B2) An initial guess of $\boldsymbol{z}_* = \boldsymbol{z}_k$ for the subproblem $\mu = \mu_{k+1}$ is asymptotically infeasible if an exact full Newton update of the primal interior point method is used. More precisely, if $\delta\rho^0_{k+1}$ is the calculated Newton update for ρ at the first iteration of the Newton solver at $\mu = \mu_{k+1}$, then as $\mu \to 0$, we see that $0 \leq \rho_k + \delta\rho^0_{k+1} \leq 1$ a.e. does not hold [24, Sec. 4.3.3].

Typically, to avoid the poor numerical behavior of (B1) and (B2), the DTO primal interior point method is reformulated as a primal-dual interior point method, eliminating the rational expressions. Since the problem is first discretized, the slack variables associated with box constraints are associated to the primal variable component-wise. This manifests as a block identity matrix within the full Hessian. The Hessian can then be reduced and the primal-dual approach is reformulated into a condensed form.

It is well known that PDE-constrained optimization solvers suffer from mesh-dependence when they do not properly treat the structure of the underlying infinite-dimensional problem [50]. In order to obtain accurate solutions, where it is clear if the material distribution indicates material or void, we may require several refinements of the mesh; in this context, it is clear that mesh-dependence would be particularly disadvantageous. The mesh-independence of our algorithm will be carefully studied in the subsequent numerical examples, and analyzed in future work.

In order to properly treat the structure of the underlying infinite-dimensional problem, we opt for an optimize-then-discretize (OTD) method. The full Hessian arising from an OTD primal-dual interior point method is no longer easily reduced, since the block associated with the slack variables is now a mass matrix, rather than the identity. To avoid solving uncondensed large systems involving three times the number of degrees of freedom of a primal approach, the goal is to develop an OTD barrier method that avoids the poor numerical behavior of (B1) and (B2). In a novel approach, we achieve this by solving the subproblems arising from the first order conditions of the enlarged feasible-set barrier functional $L^{\epsilon_{\log}}_\mu$, while still enforcing the true box constraints, $0 \leq \rho \leq 1$ a.e., with a primal-dual active set solver. Whereas in a standard barrier method, the barrier terms act as a replacement for the box constraints on ρ, *here we retain the box constraints to be handled by the primal-dual active set solver*. The barrier terms are instead used for continuation of the problem, to aid global convergence and to search for other branches of solutions.

The two inner solvers we consider are Hintermüller et al.'s *primal-dual active set strategy* (HIK) [28] and Benson and Munson's *active-set reduced space strategy* (BM) [9]. We briefly illustrate the basic approach taken to solve the individual subproblems using the log-barrier approach coupled with a primal-dual active set solver. Let $J : \mathbb{R}^n \to \mathbb{R}$ be a twice-continuously differentiable function and consider the following box-constrained nonlinear program:

$$\min_{z\in\mathbb{R}^n} J(z) \text{ subject to } a \le z \le b. \tag{3.1}$$

Here, we assume that $a, b \in \mathbb{R}^n$ such that $a < b$ (in each component) and we understand the inequality constraints $a \le z \le b$ component-wise. Next, we formulate an 'outer approximation' of (3.1) using enlarged feasible-set log-barrier terms (for any $\mu, \epsilon_{\log} > 0$):

$$\min_{z\in\mathbb{R}^n} \left\{ J(z) - \mu \sum_{i=1}^{n} [\log(z_i - (a_i - \epsilon_{\log})) + \log((b_i + \epsilon_{\log}) - z_i)] : a \le z \le b \right\}.$$

We emphasize that there are two pairs of box constraints: the true box constraints $[a, b]$ and the enlarged feasible-set box constraints $[a - \epsilon_{\log}, b + \epsilon_{\log}]$, $\epsilon_{\log} > 0$. For any fixed $\mu > 0$, the associated KKT-system has the form

$$F(z) - \lambda^a + \lambda^b = 0, \tag{3.2}$$

$$\lambda^a,\ \lambda^b \ge 0, \tag{3.3}$$

$$z - a \ge 0,\ b - z \ge 0, \tag{3.4}$$

$$\langle \lambda^a, z - a\rangle_{(\mathbb{R}^n)^*,\mathbb{R}^n} = \langle \lambda^b, b - z\rangle_{(\mathbb{R}^n)^*,\mathbb{R}^n} = 0, \tag{3.5}$$

where, $\lambda^a, \lambda^b \in (\mathbb{R}^n)^*$ are Lagrange multipliers associated with the true box constraints and

$$F(z) := J'(z) - \frac{\mu}{z - (a - \epsilon_{\log})} + \frac{\mu}{b + \epsilon_{\log} - z}, \tag{3.6}$$

where the rational expressions are interpreted component-wise. The equivalent mixed complementarity problem is given by

$$\text{either } a_i < z_i < b_i \text{ and } F(z)_i = 0, \tag{3.7}$$

$$\text{or } a_i = z_i \text{ and } F(z)_i \ge 0, \tag{3.8}$$

$$\text{or } z_i = b_i \text{ and } F(z)_i \le 0. \tag{3.9}$$

Consider the natural residual function $\varphi(x, y) = x - (x - y)_+$ where $(\cdot)_+ := \max(\cdot, 0)$. This is an example of an NCP function, a class of functions that for $x, y \in \mathbb{R}$ satisfy

$$\varphi(x, y) = 0 \text{ if and only if } x, y \ge 0,\ \ xy = 0. \tag{3.10}$$

Using φ, we note that (3.2)–(3.5) can be reformulated as the following:

$$F(z) - \lambda^a + \lambda^b = 0, \tag{3.11}$$

$$\varphi(\lambda^a, z - a) = \lambda^a - (\lambda^a - (z - a))_+ = 0, \tag{3.12}$$

$$\varphi(\lambda^b, b - z) = \lambda^b - (\lambda^b - (b - z))_+ = 0. \tag{3.13}$$

Assuming we are given a strictly enlarged-set feasible iterate $z \in \mathbb{R}^n$, $a - \epsilon_{\log} < z < b + \epsilon_{\log}$, we linearize around the point $(z, \lambda^a, \lambda^b)$ using the associated Newton-derivative and reduce the system based on the estimates of the active and inactive sets predicted by the semismooth Newton step.

In HIK, the linearized system in the direction of $(\delta z, \delta\lambda^a, \delta\lambda^b)$ is given by

$$F'(z)\delta z - \delta\lambda^a + \delta\lambda^b = -F(z) + \lambda^a - \lambda^b, \tag{3.14}$$

where $F'(z) \in \mathbb{R}^{n\times n}$ denotes the Fréchet derivative of F and

$$z_i + \delta z_i = a_i \qquad \text{if } i \in \mathfrak{A}^a = \{i : \lambda_i^a - z_i + a_i > 0\}, \tag{3.15}$$

$$z_i + \delta z_i = b_i \qquad \text{if } i \in \mathfrak{A}^b = \{i : \lambda_i^b - b_i + z_i > 0\}, \tag{3.16}$$

$$\lambda_i^a + \delta\lambda_i^a = 0 \qquad \text{if } i \in \mathfrak{I}^a = \{i : \lambda_i^a - z_i + a_i \le 0\}, \tag{3.17}$$

$$\lambda_i^b + \delta\lambda_i^b = 0 \qquad \text{if } i \in \mathfrak{I}^b = \{i : \lambda_i^b - b_i + z_i \le 0\}. \tag{3.18}$$

We define the active set by $\mathfrak{A} = \mathfrak{A}^a \cup \mathfrak{A}^b$ and the inactive set by $\mathfrak{I} = \mathfrak{I}^a \cap \mathfrak{I}^b$. By substituting (3.15)–(3.18) into (3.14) and removing the rows associated with the active set, we observe that

$$F'(z)_{\mathfrak{I},\mathfrak{I}}\delta z_{\mathfrak{I}} = -F'(z)_{\mathfrak{I},\mathfrak{A}}\delta z_{\mathfrak{A}} - F(z)_{\mathfrak{I}}. \tag{3.19}$$

We can therefore solve the reduced linear system (3.19) to find the remaining unknown components of δz.

BM attempts to solve (3.7)–(3.9) as follows. Given a feasible iterate z with respect to the true box constraints, $a \le z \le b$, the active set is defined by

$$\mathcal{A} = \{i : z_i = a_i \text{ and } F(z)_i > 0\} \cup \{i : z_i = b_i \text{ and } F(z)_i < 0\}, \tag{3.20}$$

and the inactive set is given by $\mathcal{I} = \{i\}_{i=1}^n \backslash \mathcal{A}$. The linearized system in the direction of δz takes the form

$$F'(z)_{\mathcal{I},\mathcal{I}}\delta z_{\mathcal{I}} = -F(z)_{\mathcal{I}} \text{ and } \delta z_{\mathcal{A}} = 0. \tag{3.21}$$

The next iterate is then given by $\pi(z+\delta z)$, where π is the component-wise projection onto the true box constraints, i.e.

$$\pi(z+\delta z)_i = \begin{cases} a_i & \text{if } z_i + \delta z_i < a_i, \\ z_i + \delta z_i & \text{if } a_i \le z_i + \delta z_i \le b_i, \\ b_i & \text{if } z_i + \delta z_i > b_i. \end{cases} \tag{3.22}$$

The HIK solver is a well-established method and under suitable assumptions is equivalent to a semismooth Newton method [43, 44, 56] in both finite and infinite-dimensions [28]. This equivalence ensures local superlinear convergence and under further assumptions guarantees mesh-independence [29]. Until now, the BM solver had no supporting theoretical results, although is conveniently included in PETSc [6]. Experimentally, we observe that the BM solver enjoys superlinear convergence. At first glance, the two solvers may appear quite different, but in Appendix A we prove that for a linear elliptic control problem, if the active and inactive sets coincide between the two algorithms, then the updates given by HIK and BM are identical.

One common critique of barrier methods is that the step size rules for the update of the distributed control go to zero. We observe this in numerical examples if we

use a Newton solver; however, this issue is averted when using HIK or BM. A step size of one is always taken for the update of the primal variable's active set, whereas a linesearch can be used for the update of the primal variable's inactive set. Hence, areas of the domain where the control attains the constraint do not influence the step sizes of the updates for sections of the control which are strictly feasible.

Both HIK and BM perform a pointwise projection on the iterates generated by the subproblems of the barrier functional. In the context of a classical OTD primal-dual interior point method applied to a PDE-constrained optimal control problem, under certain assumptions, Ulbrich and Ulbrich [57, 58] prove that local superlinear convergence holds if the iterates of the control and its associated Lagrange multipliers are pointwise projected to a controlled neighborhood of the central path. Although not all their assumptions hold in our case (in particular these problems are not convex), the combination of a primal-dual active set solver and barrier method mimics the computation of a Newton step of a primal-dual approach and then performing a pointwise projection. An advantage of our method is that our pointwise projection is unique and cheap to compute.

Numerically, this method only requires solving linear systems that are less than or equal to the size of the linear systems in a standard barrier method. Moreover, in the BM solver, the constrained variables can never reach the bounds of the enlarged feasible-set, ensuring the Hessian remains bounded. Furthermore, both the BM and HIK solvers remove the rows and columns in the Hessian associated with the active constraints. It is these active constraints which are the source of the unbounded eigenvalues that cause the ill-conditioning of the barrier method as μ approaches zero. In Figure 7 we give an example demonstrating that the condition number is controlled by the elimination of the active set. Removing rows and columns associated with the active set mimics the principle of Nash et al.'s *stabilized barrier method* [41, 42].

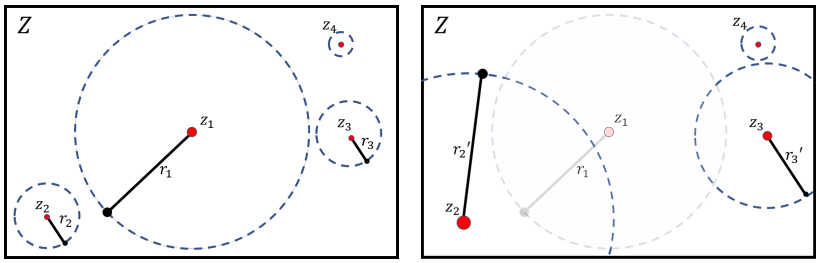
3.2. Deflation. Deflation is an algorithm for the calculation of *multiple* solutions of systems of nonlinear equations from the same initial guess. Let V and W be Banach spaces. Suppose a system of PDEs, $F(z) = 0$, $F : V \to W$ has multiple solutions $z = z_1, \dots z_n$, that we wish to find. We find the first solution by utilizing a Newton-like algorithm to find z_1. Now instead of using a standard multistart approach which may converge to the same solution, we instead introduce a modified system $G(z) = 0$ such that:

1. $G(z) = 0$ if and only if $F(z) = 0$ for $z \neq z_1$;
2. A Newton-like solver starting from any initial guess $z_* \neq z_1$ applied to G will not converge to z_1.

This process is visualized in Figure 2. In principle, one can use the same initial guess to converge to multiple solutions. The modified system is obtained by applying a *deflation operator*, $\mathcal{M}(z; z_1) : W \to W$ to F such that:

(D1) $\mathcal{M}(z; z_1)$ is invertible for all $z \neq z_1$ in a neighborhood of z_1;
(D2) $\liminf_{z \to z_1} \|\mathcal{M}(z; z_1)F(z)\| > 0$.

(D1) ensures that the resulting system has a solution if the original problem has an unknown solution, and (D2) ensures that a Newton-like method applied to the newly deflated system does not converge as $z \to z_1$. In this work we consider the shifted deflation operator $\mathcal{M}(z; z_1) = (\|z - z_1\|_V^{-2} + 1)\mathcal{I}$, where $\mathcal{I} : W \to W$ is the identity operator [21]. In particular, in all the numerical examples discussed in section 4, deflation is implemented with respect to the material distribution, i.e. $\mathcal{M}(\boldsymbol{z}; \boldsymbol{z}_1) = (\|\rho - \rho_1\|_{L^2(\Omega)}^{-2} + 1)\mathcal{I}$, where $\boldsymbol{z} = (\boldsymbol{u}, \rho, p, p_0, \lambda)$ and $\boldsymbol{z} = (\boldsymbol{u}, \rho, \lambda)$ in fluid and compliance problems, respectively.



(a) Before deflation. (b) After the deflation of z_1.

FIG. 2. *The solutions z_1, z_2, z_3 and, z_4 are zeros of the system $F(z)$. The circles around the solutions represent the basins of attraction within which a Newton-like solver converges to that particular solution.*

Deflation can be implemented very efficiently. In particular, the conditioning of the Jacobian of the deflated system does not cause computational difficulty, since the Newton update of the discrete deflated system is expressed as a scaling of the Newton update of the original discrete undeflated system via the Sherman–Morrison formula [21, Sec. 3]. Let $F_h : V_h \to W_h$ be an approximation to F on the finite-dimensional spaces V_h and W_h. Let δz_h denote the solution of the deflated Newton system evaluated at $z_h \in V_h$, to be computed, and let δy_h denote the solution of the *undeflated* Newton system of F_h, assembled at the same current iterate z_h. Let $\mathbf{z}$, $\delta\mathbf{z}$, and $\delta\mathbf{y}$ be the discrete coefficient vectors of z_h, δz_h, and δy_h, respectively. Moreover, let $m(\mathbf{z}) = \mathcal{M}(z_h, z_{1,h})$ and denote the derivative of m with respect to $\mathbf{z}$ by $m'(\mathbf{z})$. The solution of the discrete deflated Newton system can be computed by scaling $\delta\mathbf{y}$ [21, Sec. 3]:

$$\delta\mathbf{z} = \left(1 + \frac{m^{-1}(m')^{\top}(\delta\mathbf{y})}{1 - m^{-1}(m')^{\top}(\delta\mathbf{y})}\right)\delta\mathbf{y}. \tag{3.23}$$

The same formula applies if multiple solutions have been deflated, i.e. if $m(\mathbf{z}) = \mathcal{M}(z_h, z_{1,h}) \cdots \mathcal{M}(z_h, z_{n,h})$ for $n > 1$. The simple structure of (3.23) arises because the deflated residual is a (nonlinear) scalar multiple of the original residual.

In summary, in order to compute the update $\delta\mathbf{z}$ for the discretized deflated system, only the original, discretized, undeflated system is solved. Its solution $\delta\mathbf{y}$ is then scaled as in (3.23).

Deflation was first introduced in the context of polynomials by Wilkinson [61]. It was then extended to differentiable finite-dimensional maps $F : \mathbb{R}^n \to \mathbb{R}^n$ by Brown and Gearhart [13]. More recently, Farrell et al. extended the original Brown and Gearhart technique to Fréchet-differentiable maps between Banach spaces [21]. Deflation has been used to discover multiple solutions of cholesteric liquid crystals, Bose–Einstein condensates, mechanical metamaterials, aircraft stiffeners, and other applications [16, 19, 38, 46, 62]. It has also been extended to semismooth mappings [22], which is necessary in the current context of topology optimization.

3.3. Implementation of the deflated barrier method. The essential idea is to use deflation to attempt to find other branches during the continuation of the barrier parameter, as visualized in Figure 3. As summarized in Figure 4, the deflated barrier method is divided into three phases: prediction, continuation and deflation.

Prediction: Given a solution z_{k-1} at $\mu = \mu_{k-1}$, the algorithm calculates an initial

guess for the corresponding solution at $\mu = \mu_k < \mu_{k-1}$. This is done via a feasible tangent prediction method (as described in Appendix B), a classical tangent prediction method [51, Sec. 4.4.1] or a secant prediction method [51, Sec. 4.4.2]. A feasible tangent prediction method is identical to its classical counterpart but with box constraints on the predictor step to ensure the initial guess is feasible.

Continuation: Given an initial guess for each branch at the new barrier parameter μ_k, the algorithm calculates the new solution along each branch with a primal-dual active set solver whilst deflating away all solutions already known at $\mu = \mu_k$.

Deflation: At some subset of the continuation steps, the algorithm searches for new branches at $\mu = \mu_k$ using solutions on different branches found at $\mu = \mu_{k-1}$ as initial guesses. The search terminates when all the initial guesses have been exhausted (reached a maximum number of iterations without converging) or when a certain number of branches $\beta_{\max}$ have been found.

We now explain the notation used in Algorithm 3.1. Let $\boldsymbol{z} = (\boldsymbol{u}, \rho, p, p_0, \lambda)$ in the Borrvall–Petersson case and $\boldsymbol{z} = (\boldsymbol{u}, \rho, \lambda)$ in the compliance case. The value of the barrier parameter at subproblem iteration k is denoted μ_k. The initial guess for the density is denoted ρ_0 and the initial guess for the volume constraint Lagrange multiplier is denoted λ_0. The generator for the next value of μ is denoted by Θ. The μ-update can be adaptive or chosen a priori, provided it gives a strictly decreasing sequence. Under suitable conditions, the first order conditions of $L_\mu^{\epsilon_{\log}}(\boldsymbol{z})$ together with the box constraints on ρ can be reformulated into perturbed KKT conditions [58, Rem. 3] which in turn can be reformulated as a semismooth system of partial differential equations, $F_\mu(\boldsymbol{z})$. Let

$$\boldsymbol{y} = \begin{cases} (\boldsymbol{u}, p, p_0) & \text{in the Borrvall–Petersson case,} \\ \boldsymbol{u} & \text{in the compliance case.} \end{cases} \tag{3.24}$$

Let $'|_{\boldsymbol{z}_i}$ denote the Fréchet derivative with respect to $\boldsymbol{z}_i$. Let $\mathcal{S}_{\mu_k}$ denote the set of solutions, $\{\boldsymbol{z}\}_i$, found at μ_k. Let $\mathcal{M}(\cdot)$ denote the deflation operator and Z denote the function space of $\boldsymbol{z}$.

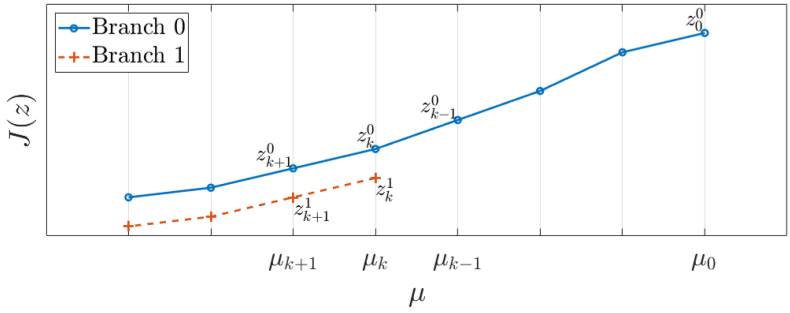


FIG. 3. *A visualization of the deflated barrier method. Branch 0 is discovered at μ_0. A predictor-corrector scheme is used to to follow the branch as μ decreases, denoted by circles. At $\mu = \mu_k$, deflation is used to discover a new solution on a different branch (branch 1), using the solution on branch 0 at $\mu = \mu_{k-1}$ as an initial guess. This newly discovered branch is then also continued as μ decreases, and is denoted by the crosses.*

Algorithm 3.1 Deflated barrier algorithm

Initialize:
  $k \leftarrow 0$ ▷ Initial iteration number
  μ_0 ▷ Initial barrier parameter
  tol ▷ Approximate solve tolerance
  $\beta_{\max}$ ▷ Maximum number of branches sought
  $\rho_0(x) \leftarrow \gamma$ ▷ Constant initial material distribution
  λ_0 ▷ Initial volume constraint multiplier
Approximately solve $(L^{\epsilon_{\log}}_{\mu_0})'|_{\boldsymbol{y}}(\boldsymbol{y}, \rho_0) = 0$. ▷ Solve state equation for $\boldsymbol{y}$
$\boldsymbol{z}_* \leftarrow (\boldsymbol{y}, \rho_0, \lambda_0)$ ▷ Initial guess
Approximately solve $F_{\mu_0}(\boldsymbol{z}) = 0$ with initial guess $\boldsymbol{z}_*$.
$S_{\mu_0} \leftarrow S_{\mu_0} \cup \{\boldsymbol{z}\}$ ▷ Include solution in solution set
$\mu_1 \leftarrow \Theta(\mu_0)$, $k \leftarrow 1$ ▷ Update μ and k
while $\mu_k \geq 0$ and $|\mathcal{S}_{\mu_{k-1}}| \neq \varnothing$ **do**
  for $\boldsymbol{z}_i \in \mathcal{S}_{\mu_{k-1}}$ **do**
    ▷ **Prediction**
    Predict solution at μ_k, denoted $\boldsymbol{z}_*$.
    ▷ **Continuation**
    Attempt to solve $\mathcal{M}\left(\mathcal{S}_{\mu_k}\right) F_{\mu_k}(\boldsymbol{z}) = 0$ with initial guess $\boldsymbol{z}_*$.
    if $\|F_{\mu_k}(\boldsymbol{z})\|_{Z^*} \leq \text{tol}$ **then**
      Solve has succeeded; set $S_{\mu_k} \leftarrow S_{\mu_k} \cup \{\boldsymbol{z}\}$.
    end if
  end for
  ▷ **Deflation**
  for $\boldsymbol{z}_j \in \mathcal{S}_{\mu_{k-1}}$ **do**
    if $|\mathcal{S}_{\mu_k}| \geq \beta_{\max}$ **then**
      break
    end if
    Attempt to solve $\mathcal{M}\left(\mathcal{S}_{\mu_k}\right) F_{\mu_k}(\boldsymbol{z}) = 0$ with initial guess $\boldsymbol{z}_j$.
    if $\|F_{\mu_k}(\boldsymbol{z})\|_{Z^*} \leq \text{tol}$ **then**
      Solve has succeeded; set $S_{\mu_k} \leftarrow S_{\mu_k} \cup \{\boldsymbol{z}\}$.
    end if
  end for
  $\mu_{k+1} \leftarrow \Theta(\mu_k)$ ▷ Choose new value of μ
  $k \leftarrow k + 1$
end while

4. Numerical results. In all examples the systems were discretized with the finite element method using FEniCS [36] and the resulting linear systems were solved by a sparse LU factorization with MUMPS [5] and PETSc [6]. The meshes were either created in FEniCS or Gmsh [26]. We present three different examples of the minimization of the power dissipation of a fluid constrained by the Stokes equations, one constrained by the Navier–Stokes equations, and two examples of the minimization of the compliance constrained by linear elasticity. Throughout the numerical examples, $h_{\min}$ denotes the minimum diameter of all simplices in the mesh, where the simplex diameter is defined as the maximum edge length. Similarly $h_{\max}$ denotes the maximum diameter of all simplices in the mesh. All solutions depicted are presented as computed by the deflated barrier method, with no truncation or postprocessing of the

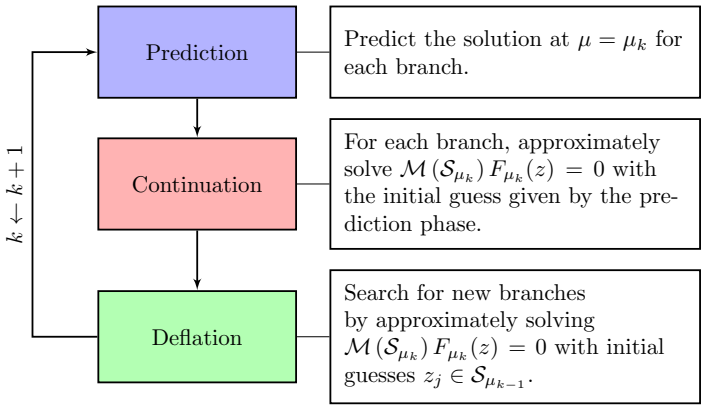


FIG. 4. *A flowchart depicting the three phases involved in the deflated barrier method.*

material distribution.

4.1. Borrvall–Petersson double-pipe. We consider the double-pipe problem with volume fraction $\gamma = 1/3$, two prescribed flow inputs and two prescribed outputs, and the boundary conditions as prescribed in Figure 5. We use α as given in (2.1), with $\overline{\alpha} = 2.5 \times 10^4$ and $q = 1/10$. Here q is a penalty parameter which controls the level of intermediate values (between zero or one) in the optimal design.

We use a Taylor–Hood $(\mathrm{CG}_2)^2 \times \mathrm{CG}_1$ finite element discretization for the velocity and pressure and CG_1 elements for the material distribution. For BM, we begin with $\mu_0 = 100$ and apply deflation immediately to find the second branch of solutions. For HIK, this strategy did not converge to the second branch, although the second branch is discovered with $\mu_0 = 105$. In both cases tangent prediction is used, as well as a damped l^2-minimizing linesearch [14, Alg. 2]. Figure 6 shows the minimizers of the double-pipe problem computed using the deflated barrier method.

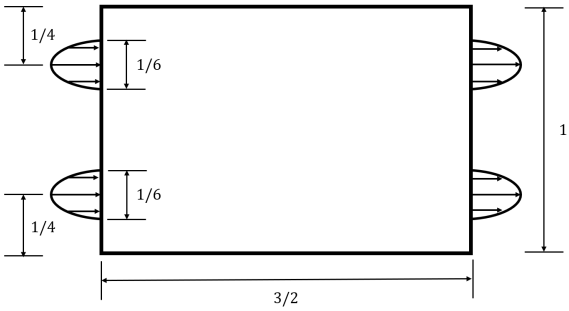


FIG. 5. *Setup of the double-pipe problem. In our tests we pick* $\boldsymbol{f} = (0,0)^\top$ *and* $\nu = 1$. *The Dirichlet boundary conditions on the velocity are* $\boldsymbol{u} = \big(1 - 144(y - 3/4)^2, 0\big)^\top$ *for the top input and output boundary flows,* $\boldsymbol{u} = \big(1 - 144(y - 1/4)^2, 0\big)^\top$ *for the bottom input and output boundary flows and* $\boldsymbol{u} = (0,0)^\top$ *everywhere else.*

In Table 1 we explore the mesh-independence of primal-dual active set solver iterations. We observe that with each refinement of the mesh, the number of iterations stay roughly constant. In particular, we notice that the behavior is consistent for both HIK and BM. This is a recurring theme and holds in subsequent examples. To

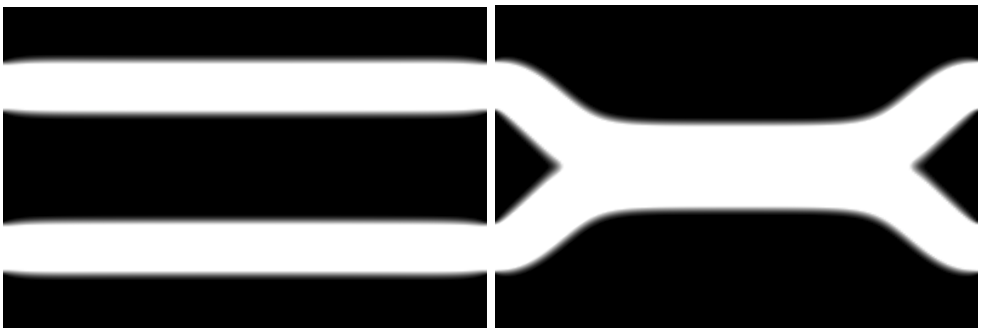


FIG. 6. *The material distribution of the local (left) and global (right) minimizer of the double-pipe optimization problem with mesh size* $h = 0.0141$. *Black corresponds to a value of* $\rho = 0$ *and white corresponds to a value of* $\rho = 1$. *The objective functional values are* $J = 32.58$ *(left) and* $J = 23.87$ *(right).*

exemplify that the mesh-independence is not an artifact of our choice of finite element spaces, we also display the results of a divergence-free Scott–Vogelius $(\mathrm{CG}_2)^2 \times \mathrm{DG}_1$ finite element discretization for the velocity and pressure and CG_1 for the material distribution. Stability of this discretization is ensured by using a barycentrically-refined mesh [45].

In Figure 7 we plot the condition number of the Hessian as in a classical barrier method, and the condition number of the Hessian with the rows and columns associated with the active-set removed. We observe that the condition number of the latter is significantly smaller, accounting for why our proposed methodology does not suffer from ill-conditioning.

BM Solver	Taylor–Hood		Branch 0			Branch 1	
h	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0283	38,256	124	0	22	115	30	22
0.0177	97,206	123	0	22	109	30	22
0.0141	151,506	110	0	22	116	29	22
HIK solver	Taylor–Hood		Branch 0			Branch 1	
h	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0283	38,256	174	0	43	261	14	43
0.0177	97,206	189	0	43	223	13	43
0.0141	151,506	173	0	43	197	13	43
BM solver	Scott–Vogelius		Branch 0			Branch 1	
$h_{\min}/h_{\max}$	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0278/0.0501	58,685	155	0	22	139	29	22
0.0139/0.0250	234,005	124	0	22	120	29	22

TABLE 1

The cumulative total numbers of primal-dual active-set solver iterations required in the continuation, deflation and prediction phases of the double-pipe problem. Branch 0 discovers the local minimum shown in Figure 6 and branch 1 discovers the global minimum. As we can see, the numbers of iterations stay roughly constant for both solvers as we refine the mesh.

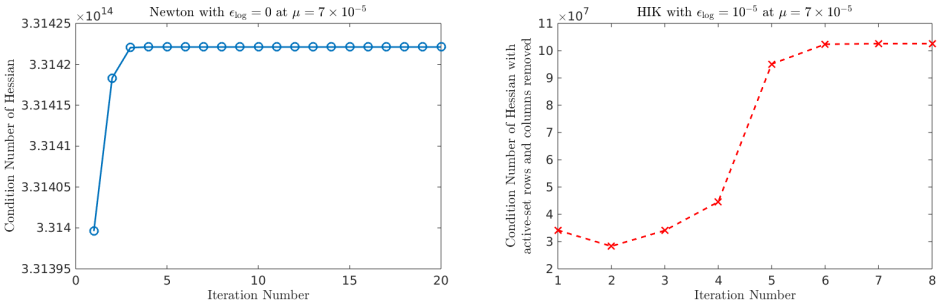


FIG. 7. *The condition number of the Hessian at each iteration of the solver in the subproblem with $\mu = 7 \times 10^{-5}$. The condition number of the Hessian of L^0_μ arising in the linear systems of a standard Newton solver (left) is six to seven orders of magnitude larger than the condition number of the Hessian of $L^{\epsilon_{log}}_\mu$ arising in the linear systems of the HIK solver (right).*

4.2. Neumann-outlet double-pipe. One could argue that fixing the outlet flows is inherently nonphysical and a more realistic model would prescribe natural boundary conditions on the outlets (while keeping the Dirichlet boundary conditions on the inlets) [17]. The correct choice of Neumann boundary conditions is nontrivial. Heywood et al. [27] provide an investigation into various formulations. We opt for the natural boundary condition,

$$(-p\mathbb{I} + 2\nu\varepsilon(\boldsymbol{u}))\,\boldsymbol{n} = \boldsymbol{0} \text{ on } \Gamma_N, \tag{4.1}$$

where $\varepsilon(\boldsymbol{u}) := (\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top)/2$ denotes the symmetrized gradient, $\mathbb{I}$ denotes the $d \times d$ identity matrix and $\Gamma_N \subset \partial\Omega$ denotes the outlets. Heywood et al. [27] note that such a formulation does not support Poiseuille flow. However, Limache et al. [34] proved that (4.1) does satisfy the principle of objectivity, which is often violated by other common formulations, including $(-p\mathbb{I} + \nu\nabla\boldsymbol{u})\,\boldsymbol{n} = \boldsymbol{0}$. The natural boundary condition (4.1) is achieved by altering the objective functional to

$$J_N(\boldsymbol{u}, \rho) = \frac{1}{2}\int_\Omega \alpha(\rho)|\boldsymbol{u}|^2 + 2\nu|\varepsilon(\boldsymbol{u})|^2 \,\mathrm{d}x. \tag{4.2}$$

Since $\mathrm{div}((\nabla\boldsymbol{u})^\top) = \nabla(\mathrm{div}(\boldsymbol{u}))$ and $\mathrm{div}(\boldsymbol{u}) = 0$, we note that the minimizers of (4.2) are the same as those of the original functional, J, combined with the natural boundary conditions as described in (4.1). The other alteration in the optimization problem is the removal of the Lagrange multiplier, p_0, since the absolute pressure level is set by the outflow boundary condition.

We employ the Taylor–Hood discretization and initialize $\mu_0 = 1000$. Deflation finds the second, third and fourth branches at $\mu = 82.4$. For $h = 0.0333$, deflation discovers branch 2, then branch 1 and 3, whereas for the other mesh sizes, deflation discovers the branches in ascending order.

The removal of an imposed outlet flow has an interesting effect. The global minimizer in the shape of a double-ended wrench is now a local minimizer. Two new $\mathbb{Z}_2$-symmetric global minimizers now exist as shown in Figure 8. This is not entirely surprising. There is a cost associated with the pipe splitting and if the optimization problem does not require the flow to leave both outlets, then it is favorable for the flow to exit via one outlet, not both. This is reflected in the resulting cost.

The mesh-independence of the algorithm is investigated in Table 2. As before, mesh-independence is observed.

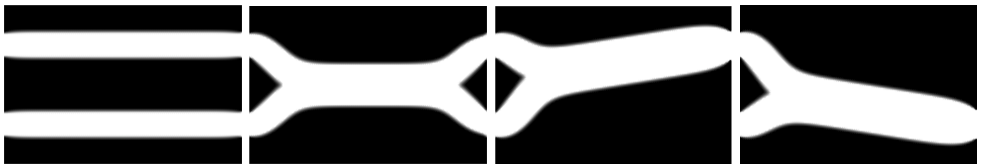


FIG. 8. *The material distribution of two local and two global minimizers of the double-pipe optimization problem with natural boundary conditions on the outlets, instead of Dirichlet conditions, with* $h = 0.0125$. *Black corresponds to a value of* $\rho = 0$ *and white corresponds to a value of* $\rho = 1$. *From left to right the objective functional values are* $J_N = 32.35$, 22.92, 18.46, *and* 18.46.

BM Solver		Branch 0			Branch 1		
h	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0333	27,455	118	0	53	108	49	34
0.0250	48,605	136	0	37	107	34	37
0.0125	193,205	113	0	35	106	45	36
		Branch 2			Branch 3		
h	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0333	27,455	166	199	55	166	149	55
0.0250	48,605	145	123	45	145	157	45
0.0125	193,205	128	151	46	128	146	46

TABLE 2

The cumulative total numbers of BM solver iterations required in the continuation, deflation and prediction phases of the double-pipe problem with natural boundary conditions on the outlets.

4.3. Roller-type pump. In this example problem [17, Sec. 2.1.4.4], the domain is given by

$$\Omega = (0,1)^2 \backslash \left\{ (x,y) \in (0,1)^2 : (x-0.5)^2 + (y-0.5)^2 \leq (0.3)^2 \right\}.$$

The boundary conditions on $\boldsymbol{u}$ are given by:

$$\boldsymbol{u} = \begin{cases} (0, 1 - 20(x-0.61)^2)^\top, & \text{if } 0.56 < x < 0.66 \text{ and } y = 0, \\ (1 - 20(y-0.95)^2, 0)^\top, & \text{if } x = 1 \text{ and } 0.9 < y < 1, \\ 10/3(y - 1/2, 1/2 - x)^\top, & \text{if } (x-0.5)^2 + (y-0.5)^2 = (0.3)^2, \\ (0,0)^\top, & \text{elsewhere.} \end{cases}$$

These boundary conditions model an inlet on the bottom of the domain and an outlet on the right of the domain with a pump rotating at a constant velocity in the center of the domain where the fluid experiences no-slip boundary conditions. We employ the Taylor–Hood discretization and initialize $\mu_0 = 1000$. Deflation finds the second branch at $\mu = 6.78$.

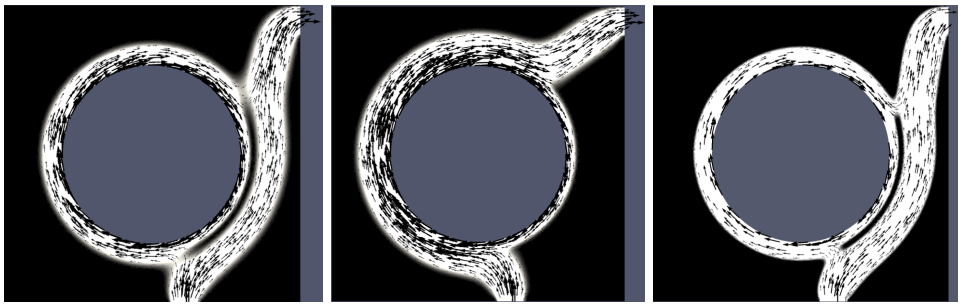
A global and local minimum of the problem are shown in Figure 9a. The local minimum chooses to avoid the pump in favor of taking the path with the shortest distance from the inlet to the outlet, while the global minimum exploits the rotation given by the pump. The local minimizer for $q = 1/10$ has areas where $\rho \approx 1/2$, which has an ambiguous physical interpretation. In order to verify whether ρ should be equal to zero or one in such areas, a mixture of grid-sequencing and continuation in q

was performed, resulting in the solution shown in Figure 9b. The mesh-independence of the algorithm is verified in Table 3.

BM solver		Branch 0			Branch 1		
$h_{\min}/h_{\max}$	Dofs	Cont.	Defl.	Pred.	Cont.	Defl.	Pred.
0.0258/0.0509	7388	260	0	55	118	80	35
0.0127/0.0255	29,174	186	0	51	75	117	25
0.0064/0.0127	113,096	177	0	46	83	99	29

TABLE 3

The cumulative total numbers of BM solver iterations required in the continuation, deflation and prediction phases of the roller-type pump problem to find the solutions shown in Figure 9a. The number of iterations are mesh-independent.



(a) The local (left) and global (right) minimizers, ρ. (b) Refined local minimizer.

FIG. 9. *(a) The material distribution of the local and global minimizers of the roller-type pump optimization problem, with $h_{min} = 6.4 \times 10^{-3}$. Black corresponds to a value of $\rho = 0$ and white corresponds to a value of $\rho = 1$. The gray area is the hole removed from the domain. The arrows indicate the direction and magnitude of the velocity, $\boldsymbol{u}$. The values of the objective functional are $J = 26.84$ (left) and $J = 22.67$ (right). (b) A mixture of grid-sequencing of the mesh where $\rho \approx 1/2$ and the continuation of q to larger values was performed on the local minimum of the roller-type pump optimization problem in order to remove areas where $\rho \approx 1/2$. The resulting refined solution has clearly defined areas of $\rho = 0$ and $\rho = 1$. Here $h_{min} = 0.0033$, $q = 0.65$ and $J = 29.17$.*

4.4. Five-holes double-pipe with Navier–Stokes. We consider the original Borrvall–Petersson double-pipe problem with Dirichlet outflow conditions, but modify the domain to include five small decagonal holes with inscribed radius 0.05 positioned at $(1/2, 1/3)$, $(1/2, 2/3)$, $(1, 1/4)$, $(1, 1/2)$ and $(1, 3/4)$, as shown in Figure 10. We further show the flexibility of our method by considering fluid flow constrained by the incompressible Navier–Stokes equations. This is achieved by introducing Lagrange multipliers, $\boldsymbol{u}_a \in H_0^1(\Omega)^d$, $p_a \in L_0^2(\Omega)$, and $p_{a,0} \in \mathbb{R}$, to enforce the Navier–Stokes equations. We then define the Lagrangian as

$$
\begin{aligned}
&L(\boldsymbol{u}, \rho, \boldsymbol{u}_a, p, p_a, p_0, p_{a,0}, \lambda) \\
&\quad = J(\boldsymbol{u}, \rho) - \int_\Omega p \operatorname{div}(\boldsymbol{u}) \mathrm{d}x - \int_\Omega p_0 p \, \mathrm{d}x - \int_\Omega \lambda(\gamma - \rho) \mathrm{d}x - \int_\Omega p_{a,0} p_a \, \mathrm{d}x \\
&\quad\quad - \int_\Omega \nu \nabla \boldsymbol{u} : \nabla \boldsymbol{u}_a + \delta(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} \cdot \boldsymbol{u}_a + \alpha(\rho)\boldsymbol{u} \cdot \boldsymbol{u}_a - p_a \operatorname{div}(\boldsymbol{u}_a) \, \mathrm{d}x,
\end{aligned} \tag{4.3}
$$

where δ denotes the (constant) fluid density. We choose $\nu = 1$ and $\delta = 1$, with other variables equal to those in the original double-pipe problem. We employ the Taylor–

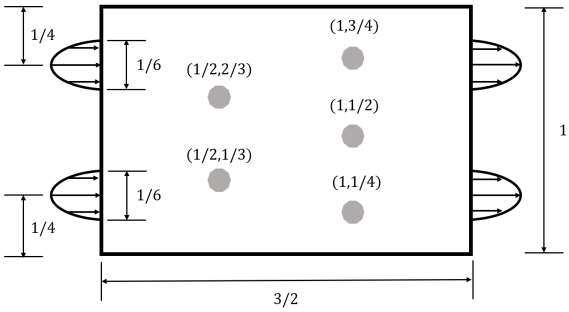


FIG. 10. *Setup of the five-holes double-pipe problem.*

Hood discretization and initialize $\mu_0 = 200$. We use feasible tangent prediction and apply an l^2-minimizing linesearch in the continuation.

The holes have the effect of substantially increasing the number of local minima, as shown in Figure 1. This example reveals that the number of local minima of a topology optimization problem is not always small and that the deflated barrier method is effective in finding many of them. A small number of solutions found exhibited regions of ambiguity $\rho \approx 1/2$, and underwent grid-sequencing and continuation in q in order to remove these areas. We note that there are more solutions that deflation did *not* find, since there are missing $\mathbb{Z}_2$ symmetric pairs which must also be solutions.

4.5. Cantilever beam. In this example we use the deflated barrier method to find multiple stationary points of compliance problems. However, due to the lack of regularity of the Lagrange multipliers associated with the box constraints on ρ, the solver exhibits mesh-dependent behavior. With each refinement of the mesh, the number of iterations required for the solver to converge increases in an unbounded way. This is difficult to resolve, and appropriate techniques to address this are the subject of ongoing research. Practically, we first run the algorithm on a coarse mesh and then use grid-sequencing to obtain refined solutions.

The two-dimensional cantilever beam optimization problem is to find minimizers of (C) that satisfy the boundary conditions

$$\begin{aligned} \sigma \boldsymbol{n} &= (0,-1)^\top && \text{on } \Gamma_N, \\ \boldsymbol{u} &= (0,0)^\top && \text{on } \Gamma_D, \\ \sigma \boldsymbol{n} &= (0,0)^\top && \text{on } \partial\Omega \backslash \{\Gamma_N \cup \Gamma_D\}, \end{aligned}$$

with domain $\Omega = (0, 1.5) \times (0, 1)$, where

$$\Gamma_D = \{(x,y) \in \partial\Omega : x = 0\},$$
$$\Gamma_N = \{(x,y) \in \partial\Omega : 0.1 \leq y \leq 0.2,\ x = 1.5\} \cup \{(x,y) \in \partial\Omega : 0.8 \leq y \leq 0.9,\ x = 1.5\}.$$

These boundary conditions describe a cantilever clamped to the y-axis with two traction forces pulling the cantilever vertically downwards in two places at $x = 1.5$. We use CG_1 finite elements for all variables. We initialize the deflated barrier method at $\mu_0 = 10$ and discover the second branch at $\mu = 4.25 \times 10^{-3}$. The two solutions found are shown in Figure 11.

4.6. Messerschmitt–Bölkow–Blohm (MBB) beam. The two-dimensional MBB beam optimization problem is to find minimizers of (C) that satisfy the bound-

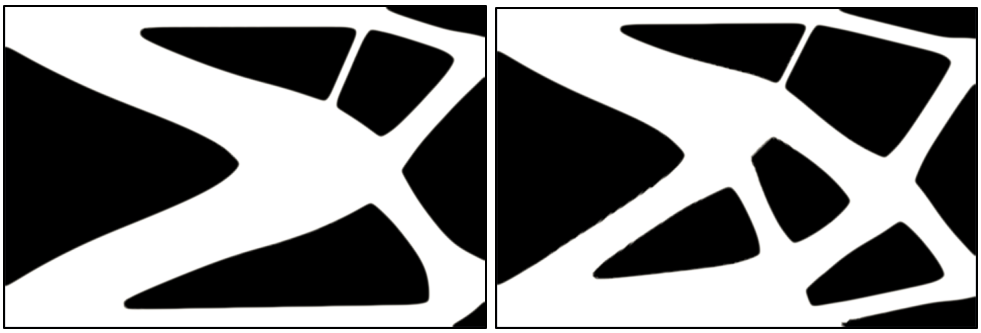


FIG. 11. *The material distribution of two solutions of the cantilever beam. The parameters are* $h_{min} = 3.56 \times 10^{-3}$, $h_{max} = 5.70 \times 10^{-2}$, $\epsilon = 4.40 \times 10^{-3}$, $\beta = 1.8 \times 10^{-4}$, $\gamma = 0.5$, $\epsilon_{SIMP} = 10^{-5}$, $p_s = 3$, *and the Lamé coefficients are* $\mu_l = 75.38$ *and* $\lambda_l = 64.62$. $J = 6.18 \times 10^{-3}$ *(left) and* $J = 6.08 \times 10^{-3}$ *(right).*

ary conditions

$$
\begin{aligned}
\boldsymbol{u} \cdot (1,0)^\top &= 0 && \text{on } \Gamma_{D_1}, \\
\boldsymbol{u} \cdot (0,1)^\top &= 0 && \text{on } \Gamma_{D_2}, \\
\sigma \boldsymbol{n} &= (0,-10)^\top && \text{on } \Gamma_N, \\
\sigma \boldsymbol{n} &= (0,0)^\top && \text{on } \partial\Omega \backslash \{\Gamma_N \cup \Gamma_{D_1} \cup \Gamma_{D_2}\},
\end{aligned}
$$

where $\Omega = (0,3) \times (0,1)$ and

$$
\begin{aligned}
\Gamma_{D_1} &= \{(x,y) \in \partial\Omega : x = 0\},\ \Gamma_{D_2} = \{(x,y) \in \partial\Omega : y = 0,\ 2.9 \le x \le 3\}, \\
\Gamma_N &= \{(x,y) \in \partial\Omega : y = 1,\ 0 \le x \le 0.1\}.
\end{aligned}
$$

These boundary conditions describe a half-beam that is fixed horizontally on the y-axis and fixed vertically at its bottom right corner on the x-axis. There is a boundary force pushing vertically downwards at the top left corner, which represents the middle of the beam when the half-beam is mirrored. We use the same finite element discretization and initialize the deflated barrier method at $\mu_0 = 50$. Deflation discovers the second branch at $\mu = 1.58 \times 10^{-1}$. As in the cantilever problem, the algorithm is mesh-dependent and grid-sequencing is used to find refinements. The two solutions found are shown in Figure 12.

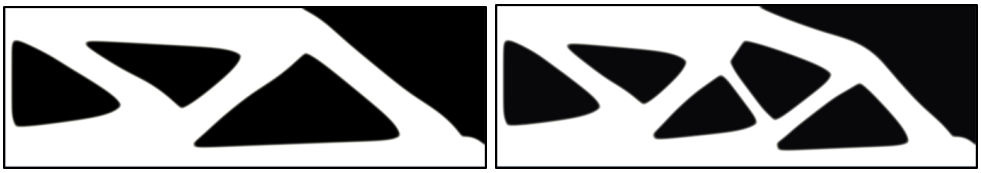


FIG. 12. *The material distribution of two solutions of the MBB beam. The parameters are* $h_{min} = 7.07 \times 10^{-3}$, $h_{max} = 2.83 \times 10^{-2}$, $\epsilon = 1.90 \times 10^{-2}$, $\beta = 9 \times 10^{-3}$, $\gamma = 0.535$, $\epsilon_{SIMP} = 10^{-5}$, $p_s = 3$, *and the Lamé coefficients are* $\mu_l = 75.38$ *and* $\lambda_l = 64.62$. $J = 0.723$ *(left) and* $J = 0.681$ *(right).*

672 **5. Conclusions.** In this work we have developed an algorithm for systemically
673 finding multiple solutions of topology optimization problems. We opted for the density
674 approach, which requires no prior knowledge of the shape or topology of the design. To
675 handle the box constraints on the material distribution ρ, we formulate an enlarged-
676 feasible set barrier functional combined with a primal-dual active set solver to ensure
677 the iterates are feasible with respect to the true box constraints. We observe com-
678 putationally that this approach does not suffer the ill-conditioning or asymptotically
679 infeasible Newton steps that normally hinder primal barrier methods. Furthermore,
680 unlike traditional primal-dual interior point methods, if the Lagrange multipliers of
681 the box constraints in the underlying continuous problem are sufficiently regular, this
682 formulation exhibits mesh-independence. The algorithm successfully found multiple
683 solutions in several problems constrained by the Stokes equations, the Navier–Stokes
684 equations, and the equations of linear elasticity.

685 **Code availability.** For reproducibility, the solver and example files to generate
686 the iteration tables and solutions can be found at https://bitbucket.org/papadopoulos/
687 deflatedbarrier/. The version of the software used in this paper is archived on Zenodo
688 [1].

689 **Appendix A. Benson and Munson's active-set reduced space solver.**
690 We show that, in the context of a linear elliptic control problem, if the active and
691 inactive sets of HIK and BM coincide, then the updates calculated for the active and
692 inactive sets are equal. In essence, we show that the algorithms produce iterates that
693 are a *half-step* out of sync, where we define the notion of a half-step below. If the
694 active and inactive sets of BM were redefined to be the same as HIK, then BM would
695 inherit the provably-good convergence properties of HIK. To our knowledge, this is the
696 first analytical result concerning BM. Although the result does not cover the nonlinear
697 case, it might help build an intuitive understanding as to why BM effectively solves
698 the semismooth formulations found in this work.
699 Consider the minimization problem

700 (A.1) $$\min_{y\in L^2(\Omega)} J(y) := \frac{1}{2}(y, Ay)_{L^2(\Omega)} - (f, y)_{L^2(\Omega)} \quad \text{subject to} \quad y \geq \phi,$$
701

702 where $(\cdot,\cdot)_{L^2(\Omega)}$ denotes the inner product in $L^2(\Omega)$, f and $\phi \in L^2(\Omega)$, and $A \in$
703 $\mathcal{L}(L^2(\Omega))$ is self-adjoint and coercive. It can be shown there exists a unique solution
704 y^* to (A.1) and there exists a Lagrange multiplier $\lambda^* \in L^2(\Omega)$ such that (y^*, λ^*) is
705 the unique solution to

706 (A.2) $$\begin{aligned} &Ay - \lambda = f, \\ &y \geq \phi,\ \lambda \geq 0,\ (\lambda, y - \phi)_{L^2(\Omega)} = 0. \end{aligned}$$
707

708 In order to avoid confusion, we denote the iterates generated by HIK by y_k and the
709 iterates generated by BM by u_k. The active and inactive sets at iteration k, $\mathfrak{A}_k$ and
710 $\mathfrak{I}_k$ in HIK and the active and inactive sets $\mathcal{A}_k$ and $\mathcal{I}_k$ in BM are defined by

711 $$\mathfrak{A}_k = \{x : \lambda_k - (y_k - \phi_i) > 0\}, \text{ and } \mathfrak{I}_k = \{x : \lambda_k - (y_k - \phi) \leq 0\},$$
712 $$\mathcal{A}_k = \{x : u_k = \phi \text{ and } F(u_k) > 0\}, \text{ and } \mathcal{I}_k = \{x : u_k > \phi \text{ or } F(u_k) \leq 0\},$$
713

714 where $F(u_k) \in L^2(\Omega)$ is the L^2-dual representation of the Fréchet derivative of $J(u_k)$.
715 As in Hintermüller et al. [28, Sec. 4], we define $E_{\mathfrak{A}_k}$ the extension-by-zero operator
716 for $L^2(\mathfrak{A}_k)$ to $L^2(\Omega)$-functions, and its adjoint $E^*_{\mathfrak{A}_k}$, the restriction operator of $L^2(\Omega)$

to $L^2(\mathfrak{A}_k)$-functions. We define $E_{\mathfrak{I}_k}$, $E^*_{\mathfrak{I}_k}$, $E_{\mathcal{A}_k}$, $E^*_{\mathcal{A}_k}$, $E_{\mathcal{I}_k}$ and $E^*_{\mathcal{I}_k}$ similarly. We note that all these restriction and extension operators are linear. We now present the infinite-dimensional description of the active-set reduced space strategy (BM).

(BM1) Choose a feasible guess $u_0 \in L^2(\Omega)$ and set $k = 0$;

(BM2) Find $\delta u_k \in L^2(\Omega)$ such that $E^*_{\mathcal{I}_k} A E_{\mathcal{I}_k} E^*_{\mathcal{I}_k} \delta u_k = -E^*_{\mathcal{I}_k}(Au_k - f)$ and $E^*_{\mathcal{A}_k} \delta u_k = 0$;

(BM3) Set $u_{k+1} = \pi(u_k + \delta u_k)$ where π is the L^2-projection onto the constraint, i.e. for any given $u \in L^2(\Omega)$, $\pi(u) \in K := \{v \in L^2(\Omega) : v \geq \phi\}$ satisfies

$$\|u - \pi(u)\|_{L^2(\Omega)} \leq \|u - v\|_{L^2(\Omega)} \text{ for all } v \in K.$$

(BM4) If convergence is reached, terminate; otherwise set $k \leftarrow k + 1$ and go to step (BM2).

THEOREM A.1. *Let y_k denote the primal variable of HIK at iteration k and let δy_k denote the update calculated at iteration k. Let λ_k denote the dual variable at iteration k. We define half steps such that the active set is updated first, i.e. $E_{\mathfrak{A}_k} y_{k+1/2} = E_{\mathfrak{A}_k} y_{k+1}$ and $E_{\mathfrak{I}_k} y_{k+1/2} = E_{\mathfrak{I}_k} y_k$.*

Let u_k denote the primal variable of BM at iteration k and let δu_k denote the update calculated at iteration k.

Suppose that $\mathcal{A}_k = \mathfrak{A}_k$, $\mathcal{I}_k = \mathfrak{I}_k$ and $E^*_{\mathfrak{I}_k} y_k = E^*_{\mathcal{I}_k} u_k$. Then the following three equalities hold;

(E1) $y_{k+1/2} = u_k$;

(E2) $E^*_{\mathfrak{I}_k} \delta y_k = E^*_{\mathcal{I}_k} \delta u_k$;

(E3) $y_{k+3/2} = u_{k+1}$.

Proof. It is shown in [28] that the update for the inactive set of HIK satisfies

$$E^*_{\mathfrak{I}_k}(A\delta y_k) = -E^*_{\mathfrak{I}_k}(Ay_k - f).$$

Expanding the left and right hand sides, we see that

$$E^*_{\mathfrak{I}_k} A E_{\mathfrak{I}_k} E^*_{\mathfrak{I}_k} \delta y_k + E^*_{\mathfrak{I}_k} A E_{\mathfrak{A}_k} E^*_{\mathfrak{A}_k} \delta y_k = -E^*_{\mathfrak{I}_k} A E_{\mathfrak{I}_k} E^*_{\mathfrak{I}_k} y_k - E^*_{\mathfrak{I}_k} A E_{\mathfrak{A}_k} E^*_{\mathfrak{A}_k} y_k + E^*_{\mathfrak{I}_k} f.$$

Subtracting the second term on the left hand side, we see that

$$E^*_{\mathfrak{I}_k} A E_{\mathfrak{I}_k} E^*_{\mathfrak{I}_k} \delta y_k = -E^*_{\mathfrak{I}_k} A E_{\mathfrak{I}_k} E^*_{\mathfrak{I}_k} y_k - E^*_{\mathfrak{I}_k} A E_{\mathfrak{A}_k} E^*_{\mathfrak{A}_k}(y_k + \delta y_k) + E^*_{\mathfrak{I}_k} f. \tag{A.3}$$

By definition $E^*_{\mathfrak{A}_k}(y + \delta y_k) = E^*_{\mathfrak{A}_k} y_{k+1/2}$ and by assumption $\mathcal{A}_k = \mathfrak{A}_k$, $\mathcal{I}_k = \mathfrak{I}_k$ and $E^*_{\mathfrak{I}_k} y_k = E^*_{\mathfrak{I}_k} u_k$. Furthermore, since by assumption $\mathcal{A}_k = \mathfrak{A}_k$ and since $E^*_{\mathfrak{A}_k} \delta y_k = E^*_{\mathfrak{A}_k}(\phi - y_k)$ as derived in [28] we observe that

$$E^*_{\mathfrak{A}_k} y_{k+1/2} = E^*_{\mathfrak{A}_k}(y_k + \phi - y_k) = E^*_{\mathfrak{A}_k} u_k. \tag{A.4}$$

Since, by definition, the first half step in HIK is only an update on the active set, we see that $E^*_{\mathfrak{I}_k} y_{k+1/2} = E^*_{\mathfrak{I}_k} y_k = E^*_{\mathfrak{I}_k} u_k$. We therefore have

$$y_{k+1/2} = u_k, \tag{A.5}$$

and (E1) holds. From (A.4), we can see that (A.3) is equivalent to

$$E^*_{\mathfrak{I}_k} A E_{\mathfrak{I}_k} E^*_{\mathfrak{I}_k} \delta y_k = -E^*_{\mathfrak{I}_k}(Au_k - f). \tag{A.6}$$

We note that (A.6) is the linear system solved to calculate the update for the inactive set of BM and hence

$$E^*_{\mathfrak{I}_k}\delta y_k = E^*_{\mathfrak{I}_k}\delta u_k. \tag{A.7}$$

Hence (E2) holds. We now show that $y_{k+3/2} = u_k$ by considering four possible cases.

(First case) Consider $C = \mathfrak{I}_k \cap \mathfrak{I}_{k+1}$. If C has measure zero, then we are done. Suppose that $|C| > 0$. Then since the dual variable is set to zero on the inactive set, we know that $E^*_C\lambda_{k+1} = 0$. Therefore, by definition of $\mathfrak{I}_{k+1}$, we know that $E^*_C y_{k+1} \geq E^*_C\phi$. Hence $E^*_C u_k + E^*_C\delta u_k \geq E^*_C\phi$ and therefore $E^*_C u_{k+1} = E^*_C\pi(u_k + \delta u_k) = E^*_C u_k + E^*_C\delta u_k = E^*_C y_{k+1}$. The first half step in HIK only changes the active set, hence $E^*_C y_{k+3/2} = E^*_C u_{k+1}$.

(Second case) Consider $C = \mathfrak{I}_k \cap \mathfrak{A}_{k+1}$. If C has measure zero, then we are done. Suppose that $|C| > 0$. Then since the dual variable is set to zero on the inactive set, we know that $E^*_C\lambda_{k+1} = 0$. Therefore, by definition of $\mathfrak{A}_{k+1}$, we know that $E^*_C y_{k+1} < E^*_C\phi$. Hence $E^*_C u_k + E^*_C\delta u_k < E^*_C\phi$ and therefore $E^*_C u_{k+1} = E^*_C\pi(u_k + \delta u_k) = E^*_C\phi$. By the half-step update of the active set, $\mathfrak{A}_{k+1}$, $E^*_C y_{k+3/2} = E^*_C\phi$. Hence $E^*_C y_{k+3/2} = E^*_C u_{k+1}$.

(Third case) Consider $C = \mathfrak{A}_k \cap \mathfrak{A}_{k+1}$. If C has measure zero, then we are done. Suppose that $|C| > 0$. This implies that $E^*_C y_{k+3/2} = E^*_C\phi$. Since $\mathfrak{A}_k = \mathcal{A}_k$, we know that $E^*_C u_{k+1} = E^*_C\phi$. Hence $E^*_C y_{k+3/2} = E^*_C u_{k+1}$.

(Fourth case) Consider $C = \mathfrak{A}_k \cap \mathfrak{I}_{k+1}$. If C has measure zero, then we are done. Suppose that $|C| > 0$. By definition of $\mathfrak{A}_k$, this implies that $E^*_C y_{k+1} = E^*_C\phi$. Furthermore, by definition of $\mathfrak{I}_{k+1}$ and since the first half step of HIK only changes the active set, we see that $E^*_C y_{k+3/2} = E^*_C\phi$. By definition of $\mathcal{A}_k$, we know that $E^*_C u_{k+1} = E^*_C\phi$. Hence $E^*_C y_{k+3/2} = E^*_C u_{k+1}$.

From the four cases, we conclude that

$$y_{k+3/2} = u_{k+1}. \tag{A.8}$$

∎

Appendix B. Feasible tangent predictor. Predictor-corrector methods are often used in tracing bifurcation diagrams [51]. The idea is that as the parameter of the problem changes, a cheap predictor generates an initial guess for the solution of the system with the new parameter. A corrector method is then used to converge from this initial guess to the true solution. In our context the primal-dual active-set solver is the corrector method. Our feasible tangent predictor method draws inspiration from the usual tangent predictor method, which solves a linear equation to find an initial guess, but applies box constraints to ensure the predicted guess is feasible.

The usual tangent predictor is derived as follows. Consider a Fréchet-differentiable equation $F(z^0, \mu^0) = 0$, where $\mu = \mu^0$ is the parameter we wish to vary. Consider a new parameter $\mu = \mu^1$ and let $\delta\mu := \mu^1 - \mu^0$. Furthermore, let $w := (z, \mu)$. The goal is to find δz such that $z^0 + \delta z \approx z^1$ where z^1 is the solution to

$$F(z^1, \mu^1) = 0. \tag{B.1}$$

A first order approximation of (B.1) is

$$0 = F(z^1, \mu^1) \approx F(z^0, \mu^0) + F'(w)\delta w = F'_z(z^0, \mu^0)\delta z + F'_\mu(z^0, \mu^0)\delta\mu. \tag{B.2}$$

Hence an initial guess, $z_* = z^0 + \delta z$, can be calculated by solving

$$F'_z(z^0, \mu^0)\delta z = -F'_\mu(z^0, \mu^0)\delta\mu, \tag{B.3}$$

for δz. In the context of the deflated barrier method this is equivalent to solving

$$(L^{\epsilon_{\log}}_{\mu^0})''|_{\boldsymbol{z},\boldsymbol{z}}(\boldsymbol{z}^0)\delta\boldsymbol{z} + (L^{\epsilon_{\log}}_{\mu^0})''|_{\boldsymbol{z},\mu}(\boldsymbol{z}^0)\delta\mu = 0, \tag{B.4}$$

for $\delta\boldsymbol{z}$. The traditional tangent predictor has no guarantee that $0 \leq \rho^0 + \delta\rho \leq 1$ a.e. To ensure that the initial guess is feasible, we instead transform (B.4) into a complementarity problem. Consider the linear operator, $T(\boldsymbol{w})$ defined by

$$\langle T(\boldsymbol{w}^0), \delta\boldsymbol{w}\rangle = (L^{\epsilon_{\log}}_{\mu^0})''|_{\boldsymbol{z},\boldsymbol{z}}(\boldsymbol{z}^0)\delta\boldsymbol{z} + (L^{\epsilon_{\log}}_{\mu^0})''|_{\boldsymbol{z},\mu}(\boldsymbol{z}^0)\delta\mu.$$

Given sufficient regularity of the dual variable $T(\boldsymbol{w})$ and the primal variable $\delta\boldsymbol{w}$, we can consider the following complementarity problem,

$$\delta\rho(x) = -\rho^0(x) \qquad \text{and } T(\boldsymbol{w}^0)(x) \geq 0, \tag{B.5}$$

$$\text{or} \quad -\rho^0(x) < \delta\rho(x) < 1 - \rho^0(x) \quad \text{and } T(\boldsymbol{w}^0)(x) = 0, \tag{B.6}$$

$$\text{or} \quad \delta\rho(x) = 1 - \rho^0(x) \qquad \text{and } T(\boldsymbol{w}^0)(x) \leq 0. \tag{B.7}$$

Solving (B.5)–(B.7) constructs a feasible tangent predictor, $\boldsymbol{z}_*$. We note that this method does not perform a pointwise projection. For example, in the topology optimization of compliance, where we require the material distribution to live in $H^1(\Omega)$, we are instead performing a H^1-projection on the prediction update. In the case where (B.6) holds a.e. in Ω, finding the feasible tangent predictor reduces to solving (B.4).

REFERENCES

[1] *Software used in 'Computing multiple solutions of topology optimization problems'*, 2020, https://doi.org/10.5281/zenodo.3710963.

[2] L. Adam, M. Hintermüller, D. Peschka, and T. M. Surowiec, *Optimization of a multi-physics problem in semiconductor laser design*, SIAM Journal on Applied Mathematics, 79 (2019), pp. 257–283, https://doi.org/10.1137/18M1179183.

[3] G. Allaire, F. Jouve, and A.-M. Toader, *A level-set method for shape optimization*, Comptes Rendus Mathematique, 334 (2002), pp. 1125–1130, https://doi.org/10.1016/S1631-073X(02)02412-3.

[4] G. Allaire, F. Jouve, and A.-M. Toader, *Structural optimization using sensitivity analysis and a level-set method*, Journal of Computational Physics, 194 (2004), pp. 363–393, https://doi.org/10.1016/j.jcp.2003.09.032.

[5] P. R. Amestoy, I. S. Duff, J.-Y. L'Excellent, and J. Koster, *A fully asynchronous multi-frontal solver using distributed dynamic scheduling*, SIAM Journal on Matrix Analysis and Applications, (2001), https://doi.org/10.1137/S0895479899358194.

[6] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkout, W. Gropp, R. Tran Mills, T. Munson, K. Rupp, P. Sana, B. Smith, S. Zampini, H. Zhang, and H. Zhang, *PETSc users manual*, Tech. Report ANL-95/11 - Revision 3.11, Argonne National Laboratory, 2019, http://www.mcs.anl.gov/petsc.

[7] M. P. Bendsøe, *Optimal shape design as a material distribution problem*, Structural Optimization, 1 (1989), pp. 193–202, https://doi.org/10.1007/BF01650949.

[8] M. P. Bendsøe and O. Sigmund, *Topology Optimization*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, https://doi.org/10.1007/978-3-662-05086-6.

[9] S. J. Benson and T. S. Munson, *Flexible complementarity solvers for large-scale applications*, Optimization Methods and Software, 21 (2003), pp. 155–168, https://doi.org/10.1080/10556780500065382.

[10] T. Borrvall, *Topology optimization of elastic continua using restriction*, Archives of Computational Methods in Engineering, 8 (2001), pp. 351–385, https://doi.org/10.1007/BF02743737.

[11] T. Borrvall and J. Petersson, *Topology optimization of fluids in Stokes flow*, International Journal for Numerical Methods in Fluids, 41 (2003), pp. 77–107, https://doi.org/10.1002/fld.426.

[12] B. Bourdin, *Filters in topology optimization*, International Journal for Numerical Methods in Engineering, 50 (2001), pp. 2143–2158, https://doi.org/10.1002/nme.116.
[13] K. M. Brown and W. B. Gearhart, *Deflation techniques for the calculation of further solutions of a nonlinear system*, Numerische Mathematik, 16 (1971), pp. 334–342, https://doi.org/10.1007/BF02165004.
[14] P. R. Brune, M. G. Knepley, B. F. Smith, and X. Tu, *Composing scalable nonlinear algebraic solvers*, SIAM Review, 57 (2015), pp. 535–565, https://doi.org/10.1137/130936725.
[15] T. E. Bruns and D. A. Tortorelli, *Topology optimization of non-linear elastic structures and compliant mechanisms*, Computer Methods in Applied Mechanics and Engineering, 190 (2001), pp. 3443–3459, https://doi.org/10.1016/S0045-7825(00)00278-4.
[16] E. G. Charalampidis, P. G. Kevrekidis, and P. E. Farrell, *Computing stationary solutions of the two-dimensional Gross-Pitaevskii equation with deflated continuation.*, Communications in Nonlinear Science and Numerical Simulation, 54 (2018), pp. 482–499, https://doi.org/10.1016/j.cnsns.2017.05.024.
[17] Y. Deng, Y. Wu, and Z. Liu, *Topology Optimization Theory for Laminar Flow: Applications in Inverse Design of Microfluidics*, Springer Singapore, Singapore, 2018, https://doi.org/10.1007/978-981-10-4687-2.
[18] Z. Doubrovski, J. C. Verlinden, and J. M. P. Geraedts, *Optimal design for additive manufacturing: opportunities and challenges*, vol. 9: 23rd International Conference on Design Theory and Methodology; 16th Design for Manufacturing and the Life Cycle Conference of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2011, pp. 635–646, https://doi.org/10.1115/DETC2011-48131.
[19] D. B. Emerson, P. E. Farrell, J. H. Adler, S. P. MacLachlan, and T. J. Atherton, *Computing equilibrium states of cholesteric liquid crystals in elliptical channels with deflation algorithms*, Liquid Crystals, 45 (2018), pp. 341–350, https://doi.org/10.1080/02678292.2017.1365385.
[20] A. Evgrafov, *State space Newton's method for topology optimization*, Computer Methods in Applied Mechanics and Engineering, 278 (2014), pp. 272–290, https://doi.org/10.1016/j.cma.2014.06.005.
[21] P. E. Farrell, Á. Birkisson, and S. W. Funke, *Deflation techniques for finding distinct solutions of nonlinear partial differential equations*, SIAM Journal on Scientific Computing, 37 (2015), pp. A2026–A2045, https://doi.org/10.1137/140984798.
[22] P. E. Farrell, M. Croci, and T. M. Surowiec, *Deflation for semismooth equations*, Optimization Methods and Software, (2019), pp. 1–24, https://doi.org/10.1080/10556788.2019.1613655.
[23] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming*, Society for Industrial and Applied Mathematics, 1990, https://doi.org/10.1137/1.9781611971316.
[24] A. Forsgren, P. E. Gill, and M. H. Wright, *Interior methods for nonlinear optimization*, SIAM Review, 44 (2002), pp. 525–597, https://doi.org/10.1137/S0036144502414942.
[25] K. R. Frisch, *The logarithmic potential method of convex programming*, tech. report, University Institute of Economics, Oslo, Norway, 1955.
[26] C. Geuzaine and J.-F. Remacle, *Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities*, International Journal for Numerical Methods in Engineering, 79 (2009), pp. 1309–1331, https://doi.org/10.1002/nme.2579.
[27] J. G. Heywood, R. Rannacher, and S. Turek, *Artificial boundaries and flux and pressure conditions for the incompressible Navier–Stokes equations*, International Journal for Numerical Methods in Fluids, 22 (1996), pp. 325–352, https://doi.org/10.1002/(SICI)1097-0363(19960315)22:5⟨325::AID-FLD307⟩3.0.CO;2-Y.
[28] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM Journal on Optimization, 13 (2003), pp. 865–888, https://doi.org/10.1137/S1052623401383558.
[29] M. Hintermüller and M. Ulbrich, *A mesh-independence result for semismooth Newton methods*, Mathematical Programming, 101 (2004), pp. 151–184, https://doi.org/10.1007/s10107-004-0540-9.
[30] R. H. Hoppe, S. I. Petrova, and V. Schulz, *Primal-dual Newton-type interior-point method for topology optimization*, Journal of Optimization Theory and Applications, 114 (2002), pp. 545–571, https://doi.org/10.1023/A:1016070928600.
[31] I. G. Jang and I. Y. Kim, *Computational study of Wolff's law with trabecular architecture in the human proximal femur using topology optimization*, Journal of Biomechanics, 41 (2008), pp. 2353–2361, https://doi.org/10.1016/j.jbiomech.2008.05.037.
[32] B. S. Lazarov and O. Sigmund, *Filters in topology optimization based on Helmholtz-type differential equations*, International Journal for Numerical Methods in Engineering, 86

(2011), pp. 765–781, https://doi.org/10.1002/nme.3072.
[33] A. V. Levy and S. Gómez, *The tunneling method applied to global optimization*, in Numerical Optimization, P. T. Boggs, ed., Society for Industrial and Applied Mathematics, 1984.
[34] A. Limache, S. Idelsohn, R. Rossi, and E. Oñate, *The violation of objectivity in Laplace formulations of the Navier–Stokes equations*, International Journal for Numerical Methods in Fluids, 54 (2007), pp. 639–664, https://doi.org/10.1002/fld.1480.
[35] 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, and A. C. To, *Current and future trends in topology optimization for additive manufacturing*, Structural and Multidisciplinary Optimization, 57 (2018), pp. 2457–2483, https://doi.org/10.1007/s00158-018-1994-3.
[36] A. Logg, K.-A. Mardal, and G. Wells, *Automated solution of differential equations by the finite element method: The FEniCS book*, Springer Science and Business Media, 84 (2012), https://doi.org/10.1007/978-3-642-23099-8.
[37] B. Maar and V. Schulz, *Interior point multigrid methods for topology optimization*, Structural and Multidisciplinary Optimization, 19 (2000), pp. 214–224, https://doi.org/10.1007/s001580050104.
[38] E. Medina, P. E. Farrell, K. Bertoldi, and C. Rycroft, *Navigating the landscape of nonlinear mechanical metamaterials for advanced programmability.*, Physical Review B, 101 (2020), https://doi.org/10.1103/PhysRevB.101.064101.
[39] H. P. Mlejnek, *Some aspects of the genesis of structures*, Structural Optimization, 5 (1992), pp. 64–69, https://doi.org/10.1007/BF01744697.
[40] L. Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Archive for Rational Mechanics and Analysis, 98 (1987), pp. 123–142, https://doi.org/10.1007/BF00251230.
[41] S. G. Nash, R. Polyak, and A. Sofer, *A numerical comparison of barrier and modified barrier methods for large-scale bound-constrained optimization*, in Large Scale Optimization, vol. 1, Springer US, Boston, MA, 1994, pp. 319–338, https://doi.org/10.1007/978-1-4613-3632-7_16.
[42] S. G. Nash and A. Sofer, *A barrier method for large-scale constrained optimization*, ORSA Journal on Computing, 5 (1993), pp. 40–53, https://doi.org/10.1287/ijoc.5.1.40.
[43] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of Operations Research, 18 (1993), pp. 227–244, https://doi.org/10.1287/moor.18.1.227.
[44] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming, 58 (1993), pp. 353–367, https://doi.org/10.1007/BF01581275.
[45] J. Qin, *On the convergence of some low order mixed finite elements for incompressible fluids*, PhD thesis, Pennsylvania State University, 1994.
[46] M. Robinson, C. Luo, P. E. Farrell, R. Erban, and A. Majumdar, *From molecular to continuum modelling of bistable liquid crystal devices*, Liquid Crystals, 44 (2017), pp. 2267–2284, https://doi.org/10.1080/02678292.2017.1290284.
[47] S. Rojas-Labanda and M. Stolpe, *Benchmarking optimization solvers for structural topology optimization*, Structural and Multidisciplinary Optimization, 52 (2015), pp. 527–547, https://doi.org/10.1007/s00158-015-1250-z.
[48] A. Schiela, *Barrier methods for optimal control problems with state constraints*, SIAM Journal on Optimization, 20 (2007), pp. 1002–1031, https://doi.org/10.1137/070692789.
[49] A. Schiela and M. Weiser, *Superlinear convergence of the control reduced interior point method for PDE constrained optimization*, Computational Optimization and Applications, 39 (2008), pp. 369–393, https://doi.org/10.1007/s10589-007-9057-5.
[50] T. Schwedes, D. A. Ham, S. W. Funke, and M. D. Piggott, *Mesh dependence in PDE-constrained optimisation*, Springer, 2017, https://doi.org/10.1007/978-3-319-59483-5_2.
[51] R. Seydel, *Practical Bifurcation and Stability Analysis*, vol. 5 of Interdisciplinary Applied Mathematics, Springer New York, New York, NY, 3 ed., 2010, https://doi.org/10.1007/978-1-4419-1740-9.
[52] O. Sigmund, *Design of material structures using topology optimization*, PhD thesis, Technical University of Denmark, 1994.
[53] J. Sokolowski and A. Zochowski, *On the topological derivative in shape optimization*, SIAM Journal on Control and Optimization, 37 (1999), pp. 1251–1272, https://doi.org/10.1137/S0363012997323230.
[54] M. Stolpe and K. Svanberg, *On the trajectories of penalization methods for topology optimization*, Structural and Multidisciplinary Optimization, 21 (2001), pp. 128–139, https://doi.org/10.1007/s001580050177.
[55] K. Svanberg, *The method of moving asymptotes - a new method for structural optimization*, International Journal for Numerical Methods in Engineering, 24 (1987), pp. 359–373, https:

//doi.org/10.1002/nme.1620240207.

[56] M. Ulbrich, *Semismooth Newton methods for operator equations in function spaces*, SIAM Journal on Optimization, 13 (2003), pp. 805–841, https://doi.org/10.1137/s1052623400371569.

[57] M. Ulbrich and S. Ulbrich, *Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds*, SIAM Journal of Control and Optimization, 38 (2000), pp. 1938–1984, https://doi.org/10.1137/S0363012997325915.

[58] M. Ulbrich and S. Ulbrich, *Primal-dual interior-point methods for PDE-constrained optimization*, Mathematical Programming, 117 (2009), pp. 435–485, https://doi.org/10.1007/s10107-007-0168-7.

[59] M. Y. Wang, X. Wang, and D. Guo, *A level set method for structural topology optimization*, Computer Methods in Applied Mechanics and Engineering, 192 (2003), pp. 227–246, https://doi.org/10.1016/S0045-7825(02)00559-5.

[60] M. Weiser, T. Gänzler, and A. Schiela, *A control reduced primal interior point method for a class of control constrained optimal control problems*, Computational Optimization and Applications, 41 (2008), pp. 127–145, https://doi.org/10.1007/s10589-007-9088-y.

[61] J. H. Wilkinson, *Rounding errors in algebraic processes*, Dover Publications, Inc., New York, 1994. Reprint of the 1963 original [Prentice-Hall, Englewood Cliffs, NJ].

[62] J. Xia, P. E. Farrell, and S. G. P. Castro, *Nonlinear bifurcation analysis of stiffener profiles via deflation techniques*, Thin Walled Structures, 149 (2020), p. 106662, https://doi.org/10.1016/j.tws.2020.106662.

[63] Y. Xie and G. Steven, *A simple evolutionary procedure for structural optimization*, Computers and Structures, 49 (1993), pp. 885–896, https://doi.org/10.1016/0045-7949(93)90035-C.

[64] S. Zhang and J. A. Norato, *Finding better local optima in topology optimization via tunneling*, in Volume 2B: 44th Design Automation Conference, vol. 17, American Society of Mechanical Engineers, 2018, p. 103, https://doi.org/10.1115/DETC2018-86116.