The Journal of Geometric Analysis (2024) 34:206
https://doi.org/10.1007/s12220-024-01663-0

Efficient and Accurate Separable Models for Discretized Material Optimization: A Continuous Perspective Based on Topological Derivatives

Peter Gangl[1,2] · **Nico Nees**[2] · **Michael Stingl**[2]

Received: 3 February 2024 / Accepted: 10 April 2024

Abstract
Multi-material design optimization problems can, after discretization, be solved by the iterative solution of simpler sub-problems which approximate the original problem at an expansion point to first order. In particular, models constructed from convex separable first order approximations have a long and successful tradition in the design optimization community and have led to powerful optimization tools like the prominently used method of moving asymptotes (MMA). In this paper, we introduce several new separable approximations to a model problem and examine them in terms of accuracy and fast evaluation. The models can, in general, be nonconvex and are based on the Sherman–Morrison–Woodbury matrix identity on the one hand, and on the mathematical concept of topological derivatives on the other hand. We show a surprising relation between two models originating from these two—at a first sight—very different concepts. Numerical experiments show a high level of accuracy for two of our proposed models while also their evaluation can be performed efficiently once enough data has been precomputed in an offline stage. Additionally it is demonstrated that suboptimal decisions can be avoided using our most accurate models.

Keywords Discretized material optimization · Separable models · Topological derivative · Sherman–Morrison–Woodbury formula

✉ Peter Gangl
peter.gangl@ricam.oeaw.ac.at

Nico Nees
nico.nees@fau.de

Michael Stingl
stingl@math.fau.de

1 Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Straße 69, Linz 4040, Austria

2 Chair of Applied Mathematics (Continuous Optimization), Friedrich-Alexander-Universität Erlangen-Nürnberg, Cauerstraße 11, Erlangen 91058, Germany

Published online: 29 April 2024

1 Introduction

The goal of computational design optimization is to find an optimal arrangement of possibly multiple materials inside a design region of a computational domain. Here, optimality is considered with respect to a given cost function, which most often depends on the solution of a constraining partial differential equation (PDE). Thus, a general PDE-constrained design optimization problem reads

$$\Omega^* = \arg\min_{\Omega} J(\Omega, u) \quad \text{subject to} \quad e(\Omega; u) = 0 \tag{1}$$

where $e(\cdot;\cdot)$ represents the PDE constraint and Ω can also be understood as a set of subdomains corresponding to different materials. There exist different classes of approaches to solving problems of this type. While shape optimization methods [1] can only modify existing boundaries or interfaces between subdomains in a smooth way, topology optimization approaches [2] can also alter the topology of a design and can thus admit more general solutions. In topology optimization, the design is most often represented by either a level set function [3, 4] or by means of a density function [5] that interpolates between different material properties. Note that both kinds of approaches can be extended to the case of multiple materials, see, e.g., [6, 7] or [8]. Typically, the constraining PDE is solved by a numerical method, most often by the finite element method. When approaching a design optimization problem of the type (1) by a gradient-based approach, one first has to decide whether the differentiation with respect to the design variable should be carried out before or after discretization of the problem. It should be noted that, depending on the chosen discretization and parametrization of the design, these two ways need not yield the same discrete sensitivities [9]. We mention that, of course, the numerical analysis in finite dimensions is only a part of the solution of shape and topology optimization problems. Such an analysis should be completemented by the results on the infinite dimensional model including the existence of an optimal shape and the convergence of derivative based optimization methods. Such results are obtained usually by using appropriate regularization techniques. We refer the reader, e.g., to [10].

In this paper we are interested in a (multi-material) topology optimization problem which we parametrize by a density function. Here, we focus on the approach where we first discretize the problem and then aim to solve the discretized, finite-dimensional problem. Given a computational domain $\mathsf{D} \subset \mathbb{R}^2$ which is discretized using a fixed structured mesh consisting of m triangular elements with n vertices, we aim at finding the optimal material distribution $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ minimizing the heat compliance,

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda} \in [\underline{\lambda}, \overline{\lambda}]^m} \mathcal{J}(\boldsymbol{\lambda}) \tag{2}$$

with $\mathcal{J}(\boldsymbol{\lambda}) := \boldsymbol{f}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{f}$. Here, $\boldsymbol{K}(\boldsymbol{\lambda}) \in \mathbb{R}^{n \times n}$ and $\boldsymbol{f} \in \mathbb{R}^n$ represent, respectively, the (invertible) stiffness matrix and the load vector corresponding to a discretization by means of piecewise linear, globally continuous finite elements where the material coefficient in the ℓ-th element T_ℓ is given by $\boldsymbol{\lambda}_\ell$, $\ell \in \{1, \ldots, m\}$. The problem may be

subject to additional constraints, e.g., on the volume of given materials, or enriched by terms that penalize the appearance of intermediate materials or that regularize the obtained designs by filtering [11].

The idea of *sequential global programming* (SGP) [12] is the following: Instead of solving an optimization problem like (2) over $\mathbb{R}^m$ directly, one solves a sequence of simpler optimization problems with cost function $\hat{\mathcal{J}}[\boldsymbol{\lambda}^{(k)}](\boldsymbol{\lambda})$ which approximate the original problem with cost function $\mathcal{J}(\boldsymbol{\lambda})$ at an expansion point $\boldsymbol{\lambda}^{(k)}$ to first order. The solution of the simpler optimization problem at iteration k is subsequently used as an expansion point $\boldsymbol{\lambda}^{(k+1)}$ in iteration $k+1$, i.e.,

$$\boldsymbol{\lambda}^{(k+1)} = \underset{\boldsymbol{\lambda}\in[\underline{\lambda},\overline{\lambda}]^m}{\arg\min}\, \hat{\mathcal{J}}[\boldsymbol{\lambda}^{(k)}](\boldsymbol{\lambda}).$$

A class of approximate models that is of particular interest is the class of *separable* models. The use of convex, separable approximations has a long tradition in design optimization, see, e.g., [13–15] and have lead to powerful software realizations like CONLIN [16] or the method of moving asymptotes [17]. Such models allow to solve the approximate optimization problem that is posed over $\mathbb{R}^m$ by solving merely m one-dimensional optimization problems. These one-dimensional sub-problems can often be solved to global optimality. This observation holds true trivially for the more traditional convex separable approximations, used, e.g., in MMA. However, for separable approximations, convexity is not a strict requirement. It is clear that the convergence speed of an SGP algorithm strongly depends on the quality of the approximating model $\hat{\mathcal{J}}[\boldsymbol{\lambda}^{(k)}]$. In this paper, we propose different first order separable models $\hat{\mathcal{J}}[\boldsymbol{\lambda}^{(k)}]$ and numerically examine them in terms of accuracy and efficiency of evaluation. But it is not only the efficiency, which is of interest. This becomes particularly evident, when topology optimization or discrete material optimization problems are studied. The usual way to deal with such problems is to use a combined relaxation and penalization scheme, see, e.g., [5] for an introduction to that topic. While such approaches are very successful in practice, in extreme cases it can happen that any feasible integer solution satisfies first order optimality conditions for the continuous relaxations. Thus, there is a certain risk that rather poor local minimizers are obtained. While in literature so-called continuation strategies provide a viable concept to cope with that situation, in this paper we demonstrate that it is in particular the approximation quality in the sub-problem, which can help to avoid 'wrong' decisions taken in the course of the iterations.

We will investigate models $\hat{\mathcal{J}}$ that exploit the Sherman–Morrison–Woodbury matrix identity on the discrete level and are thus purely algebraic. On the other hand, we will consider the mathematical concept of topological derivatives [18] which is a notion defined on a purely continuous setting. We will draw some, at a first glance, surprising connections between these two types of approaches and present some models that are at the same time accurate approximations of the original problem and cheap to evaluate.

The rest of this paper is organized as follows: In Sect. 2, we introduce the continuous model problem and its finite element discretization and recall the notions of topological derivatives, separable approximations of optimization problems and

also the Sherman–Morrison–Woodbury formula. Next, we introduce a first efficient model based on this formula that is based on a diagonal approximation of the stiffness matrix in Sect. 3. Subsequently, we introduce a different model that is motivated by the concept of topological derivatives in Sect. 4. In Sect. 5, inspired by the procedure of the previous section, we introduce another accurate and efficient to evaluate approximation to the exact Sherman–Morrison–Woodbury model. A relation between these latter two models is established in Sect. 6. Finally, we examine all introduced models numerically in Sect. 7.

2 Preliminaries

In this section, we will introduce the model problem and collect some mathematical preliminaries, which we will make use of in later sections. In particular, we introduce the mathematical concept of topological derivatives, the concept of separable first order approximations of a continuously differentiable function $f : \mathbb{R}^N \to \mathbb{R}$ and recall the Sherman–Morrison–Woodbury matrix identity.

Notation. Vector quantities and matrices will be denoted by bold-face symbols and we will denote the j-th component of a vector $\boldsymbol{v} \in \mathbb{R}^N$ by a sub-index $\boldsymbol{v}_j$. We will denote the i-th Cartesian unit vector in $\mathbb{R}^N$ by $\boldsymbol{e}^{(i)}$. The identity matrix of dimension N will be denoted by the symbol $\boldsymbol{I}_N$ and for a square matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ we denote by $\mathrm{diag}\boldsymbol{A} \in \mathbb{R}^{N\times N}$ the diagonal matrix corresponding to $\boldsymbol{A}$, i.e., $(\mathrm{diag}\boldsymbol{A})_{i,i} = \boldsymbol{A}_{i,i}$ and $(\mathrm{diag}\boldsymbol{A})_{i,j} = 0$ for $i \neq j, i, j \in \{1, \dots N\}$. We denote by $B_r(x)$ the ball of radius r centered at the point x. Moreover, given a set $A \subset \mathbb{R}^d$, we denote the characteristic function of the set A by $\chi_A(x)$, i.e., $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise. The space of square integrable functions over a domain D is denoted by $L^2(\mathsf{D})$ and the subspace of $L^2(\mathsf{D})$ functions whose weak gradient is also a $L^2(\mathsf{D})$ function is denoted by $H^1(\mathsf{D})$. Finally, given a function f defined on a domain D and a subdomain $\Omega \subset \mathsf{D}$, we will denote the restriction of f to Ω by $f|_\Omega$.

2.1 Model Problem

As a model problem, we consider a stationary heat equation on a bounded Lipschitz domain $\mathsf{D} \subset \mathbb{R}^d$. We are interested in finding the optimal material distribution within D such that the heat compliance is minimized. We first state the model problem in its continuous version before introducing the discretized problem, which we are actually interested in. In this paper, we restrict ourselves to space dimensions $d = 1$ and $d = 2$, but remark that most concepts treated here can be extended (with some effort) also to three space dimensions.

2.1.1 Continuous Model Problem for Two Materials

For the spatial dimension $d \in \{1, 2\}$, we consider the bounded Lipschitz domain $\mathsf{D} = (0, 1)^d \subset \mathbb{R}^d$ and a given heat source $f \in L^2(\mathsf{D})$. Given a polygonal set $\Omega \subset \mathsf{D}$, let the piecewise constant heat conductivity λ be defined by

$$\lambda_\Omega(x) = \begin{cases} \lambda^{\text{in}}, & x \in \Omega, \\ \lambda^{\text{out}}, & x \in \mathsf{D} \setminus \overline{\Omega}, \end{cases} \tag{3}$$

for two positive constants $\lambda^{\text{in}}, \lambda^{\text{out}} > 0$. We assume the boundary of the computational domain to be composed of a Dirichlet and a Neumann boundary, $\partial\mathsf{D} = \Gamma_D \cup \Gamma_N$ with $\Gamma_D \cap \Gamma_N = \emptyset$, where inhomogeneous Dirichlet data g_D and Neumann data g_N are prescribed, respectively. We are interested in minimizing the heat compliance,

$$J(u) := \int_{\mathsf{D}} f(x)\, u(x)\, \mathrm{d}x + \int_{\Gamma_N} g_N(x)\, u(x)\, \mathrm{d}s_x \tag{4}$$

subject to a stationary heat equation. The weak formulation of the problem at hand reads

$$\inf_{\Omega} J(u) \tag{5a}$$

$$\text{s.t. } u \in V_g : \int_{\mathsf{D}} \lambda_\Omega(x) \nabla u(x) \cdot \nabla v(x)\, \mathrm{d}x = \int_{\mathsf{D}} f(x) v(x)\, \mathrm{d}x + \int_{\Gamma_N} g_N(x) v(x)\, \mathrm{d}s_x \quad \forall v \in V_0 \tag{5b}$$

with the ansatz and test spaces

$$V_g := \{v \in H^1(\mathsf{D}) : v|_{\Gamma_D} = g_D\}, \qquad V_0 := H^1_{\Gamma_D}(\mathsf{D}) := \{v \in H^1(\mathsf{D}) : v|_{\Gamma_D} = 0\}.$$

For a given subdomain Ω and assuming that $|\Gamma_D| > 0$, due to the Lemma of Lax-Milgram, problem (5b) admits a unique solution which we denote by u_Ω. Thus, we introduce the reduced cost function $\mathcal{J}(\Omega) := J(u_\Omega)$. We assume that the solution u_Ω is sufficiently regular such that a point evaluation of its gradient $\nabla u_\Omega(z)$ is well-defined for all points $z \in D \setminus \partial\Omega$. When the set Ω is clear from the context, we will drop the index Ω and just write u instead of u_Ω. For simplicity and without loss of generality, we assume $g_D = 0$ and thus have $V_g = V_0$. The general case can be obtained by homogenization of the Dirichlet data. Note that, in the case $d = 1$, the boundary of D consists of two points, $\partial\mathsf{D} = \{0, 1\}$. Thus, integrals over $\Gamma_N \subset \partial\mathsf{D}$ have to be understood as point evaluations.

The adjoint state corresponding to optimization problem (5) is the unique solution $p \in V_0$ of

$$\int_{\mathsf{D}} \lambda_\Omega(x) \nabla v(x) \cdot \nabla p(x)\, \mathrm{d}x = -\int_{\mathsf{D}} f(x) v(x)\, \mathrm{d}x - \int_{\Gamma_N} g_N(x) v(x)\, \mathrm{d}s_x \quad \forall v \in V_0. \tag{6}$$

Thus, it can be seen that $p = -u$.

2.1.2 Discrete Model Problem

Next, we introduce the discretization of (5) by means of piecewise linear, globally continuous finite elements. For that purpose, let $\mathcal{T} = \{T_1, \dots, T_m\}$ denote a set of open simplicial elements (i.e., intervals in 1D or triangles in 2D) which form a subdivision of the computational domain D, i.e.,

$$\overline{\mathsf{D}} = \bigcup_{\ell=1}^{m} \overline{T}_\ell, \qquad T_i \cap T_j = \emptyset \text{ for } i \neq j.$$

Moreover, we assume that the subdomain Ω is resolved by the mesh $\mathcal{T}$, i.e., $\partial\Omega \cap T_\ell = \emptyset$ for all $\ell \in \{1, \dots, m\}$. Let $n \in \mathbb{N}$ denote the number of vertices in the mesh, $\{\varphi_1, \dots, \varphi_n\}$ the nodal basis and $V_h := \operatorname{span}\{\varphi_1, \dots, \varphi_n\} \cap H^1_{\Gamma_D}(\mathsf{D})$.

Let a vector of conductivity values $\boldsymbol{\lambda} \in \mathbb{R}^m$ be given. Note that we will sometimes identify a vector $\boldsymbol{\lambda} \in \mathbb{R}^m$ of material values with a piecewise constant material function $\lambda(x) := \sum_{\ell=1}^m \chi_{T_\ell}(x)\boldsymbol{\lambda}_\ell$. For given $\boldsymbol{\lambda} \in \mathbb{R}^m$, the discrete version of the boundary value problem (5b) reads

$$\boldsymbol{K}(\boldsymbol{\lambda})\boldsymbol{u} = \boldsymbol{f} \tag{7}$$

where the stiffness matrix $\boldsymbol{K}(\boldsymbol{\lambda}) \in \mathbb{R}^{n\times n}$ and the load vector $\boldsymbol{f} \in \mathbb{R}^n$ can be written as

$$\boldsymbol{K}(\boldsymbol{\lambda}) = \sum_{\ell=1}^{m} \boldsymbol{\lambda}_\ell \tilde{\boldsymbol{B}}_\ell K^{(\ell)}_{loc} \tilde{\boldsymbol{B}}_\ell^\top, \qquad \boldsymbol{f} = \sum_{\ell=1}^{m} \tilde{\boldsymbol{B}}_\ell f^{(\ell)}_{loc} \tag{8}$$

with the local stiffness matrix $K^{(\ell)}_{loc} \in \mathbb{R}^{(d+1)\times(d+1)}$ and local load vector $f^{(\ell)}_{loc} \in \mathbb{R}^{d+1}$,

$$\left(K^{(\ell)}_{loc}\right)_{i,j} = \int_{T_\ell} \nabla\varphi_{\ell,j} \cdot \nabla\varphi_{\ell,i} \,\mathrm{d}x, \qquad i, j \in \{1, \dots, d+1\},$$
$$\left(f^{(\ell)}_{loc}\right)_i = \int_{T_\ell} f\varphi_{\ell,i} \,\mathrm{d}x + \int_{\Gamma_N \cap \bar{T}_\ell} g_N \varphi_{\ell,i} \,\mathrm{d}s_x, \quad i \in \{1, \dots, d+1\},$$

and the local-to-global operator $\tilde{\boldsymbol{B}}_\ell \in \mathbb{R}^{n\times(d+1)}$ satisfying $(\tilde{\boldsymbol{B}}_\ell)_{i,j} = 1$ if i is the global index of the j-th vertex of element T_ℓ, and $(\tilde{\boldsymbol{B}}_\ell)_{i,j} = 0$ else. Here, $\varphi_{\ell,i}$, $i = 1, \dots, d+1$, denotes the i-th basis functions that has non-zero support on T_ℓ. Since we are dealing with piecewise linear and globally continuous finite elements, the local stiffness matrix can be written as

$$K^{(\ell)}_{loc} = \boldsymbol{D}_\ell \boldsymbol{D}_\ell^\top \tag{9}$$

with some constant matrices $\boldsymbol{D}_\ell \in \mathbb{R}^{(d+1)\times d}$ depending solely on the coordinates of the vertices of element T_ℓ. Thus, defining $\boldsymbol{B}_\ell := \tilde{\boldsymbol{B}}_\ell \boldsymbol{D}_\ell \in \mathbb{R}^{n\times d}$, the stiffness matrix can also be written as

$$\boldsymbol{K}(\boldsymbol{\lambda}) = \sum_{\ell=1}^{m} \lambda_\ell \boldsymbol{B}_\ell \boldsymbol{B}_\ell{}^\top. \tag{10}$$

Remark 1 In dimension $d = 1$, the matrix $\boldsymbol{D}_\ell \in \mathbb{R}^{2\times 1}$ corresponding to an element $T_\ell = (x_{\ell-1}, x_\ell)$ is given by

$$\boldsymbol{D}_\ell = \frac{1}{\sqrt{x_\ell - x_{\ell-1}}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

For $d = 2$ and a triangular element T_ℓ with vertices $\boldsymbol{x}_{\ell,1}, \boldsymbol{x}_{\ell,2}, \boldsymbol{x}_{\ell,3}$ in counter-clockwise enumeration, the matrix $\boldsymbol{D}_\ell \in \mathbb{R}^{3\times 2}$ reads

$$\boldsymbol{D}_\ell = \sqrt{\frac{1}{2}\det \boldsymbol{J}_\ell} \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{J}_\ell^{-1}$$

with $\boldsymbol{J}_\ell = \left(\boldsymbol{x}_{\ell,2} - \boldsymbol{x}_{\ell,1}\ \boldsymbol{x}_{\ell,3} - \boldsymbol{x}_{\ell,1}\right) \in \mathbb{R}^{2\times 2}$.

Finally, the Dirichlet boundary conditions on nodes $\boldsymbol{v}^{(i)}$ on $\overline{\Gamma}_D$ are incorporated by setting $(\boldsymbol{K}(\boldsymbol{\lambda}))_{i,i} = 1$ and $(\boldsymbol{K}(\boldsymbol{\lambda}))_{i,j} = (\boldsymbol{K}(\boldsymbol{\lambda}))_{j,i} = 0$ for $i \neq j$ and $\boldsymbol{f}_i = g_D(\boldsymbol{v}_i)$. Note that, for $\boldsymbol{\lambda} \in [\underline{\lambda}, \overline{\lambda}]^m$ with $\underline{\lambda} > 0$, the stiffness matrix after incorporation of the Dirichlet boundary conditions is invertible. Thus, we can define the solution vector

$$\boldsymbol{u}(\boldsymbol{\lambda}) := \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{f} \tag{11}$$

and the corresponding discrete solution $u_h(x) := \sum_{i=1}^{n} \boldsymbol{u}_i \varphi_i(x)$, and we introduce the discrete compliance function $\mathcal{J} : [\underline{\lambda}, \overline{\lambda}]^m \to \mathbb{R}$,

$$\mathcal{J}(\boldsymbol{\lambda}) := \boldsymbol{f}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{f}. \tag{12}$$

Remark 2 In order to obtain practically interesting multi-material designs, the discretized problem should additionally include a mechanism to penalize intermediate material. This can be done by adding a term of the form $J_{\text{gray}}(\boldsymbol{\lambda}) = \sum_\ell |T_\ell|(\lambda_\ell - \lambda^{\text{in}})(\lambda^{\text{out}} - \lambda_\ell)$ (or an extension of this to multiple materials) to the cost function. On the other hand, it is well-known that topology optimization problems of the type (2) often lack a solution which can be observed numerically in the form of mesh-dependent optimized designs. In order to obtain a well-defined problem, one typically introduces a length scale in the form of a filter radius. This can be realized by adding a term of the form $J_{\text{reg}}(\boldsymbol{\lambda}) = \|\mathbf{F}_R \boldsymbol{\lambda} - \boldsymbol{\lambda}\|_2^2$ to the cost function. Here, $\mathbf{F}_R$ is a filtering operator with a given length scale R. Thus, one typically is interested in minimizing an enriched cost function $\tilde{\mathcal{J}}(\boldsymbol{\lambda}) = \mathcal{J}(\boldsymbol{\lambda}) + \gamma_1 J_{\text{gray}}(\boldsymbol{\lambda}) + \gamma_2 J_{\text{reg}}(\boldsymbol{\lambda})$. Since the functionals $J_{\text{gray}}(\boldsymbol{\lambda})$ and $J_{\text{reg}}(\boldsymbol{\lambda})$ are often separable and can be evaluated efficiently by default, we will ignore these terms for the rest of this paper. For a more detailed discussion on this aspect for multi-material topology optimization, see [19, Sec. 2.2].

Later on, we will make use of the following relation.

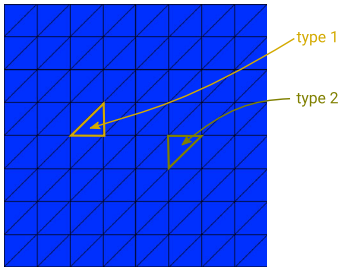


Fig. 1 Example of structured mesh as used in this paper consisting of only two different types of elements

Lemma 1 *Let u_h a piecewise linear and globally continuous finite element function on a given simplicial mesh in $\mathbb{R}^d$, $d \in \{1, 2\}$ with vector of basis coefficients $\boldsymbol{u} \in \mathbb{R}^n$ and let $\boldsymbol{B}_\ell = \tilde{\boldsymbol{B}}_\ell \boldsymbol{D}_\ell$ defined as above. Then it holds for any $\ell \in \{1, \dots, m\}$*

$$\boldsymbol{B}_\ell{}^\top \boldsymbol{u} = \sqrt{|T_\ell|}\, \nabla u_h|_{T_\ell}. \tag{13}$$

Proof First note that $\boldsymbol{u}^\top \boldsymbol{B}_\ell = \boldsymbol{u}^\top \tilde{\boldsymbol{B}}_\ell \boldsymbol{D}_\ell$ and that $\boldsymbol{u}^\top \tilde{\boldsymbol{B}}_\ell$ is the vector of local degrees of freedom on element T_ℓ. Thus, for $d = 1$, we get $\boldsymbol{u}^\top \boldsymbol{B}_\ell = [\boldsymbol{u}_{\ell,1}, \boldsymbol{u}_{\ell,2}]\boldsymbol{D}_\ell = \sqrt{|T_\ell|}(\boldsymbol{u}_{\ell,2} - \boldsymbol{u}_{\ell,1})/|T_\ell|$. The assertion follows by recalling that $|T_\ell|$ is the length of the interval T_ℓ and that u_h is linear on T_ℓ.

In order to see the relation in the case $d = 2$, let Φ_ℓ denote the affine transformation with Jacobian matrix $\boldsymbol{J}_\ell$ that maps the reference triangle with vertices $(0, 0)^\top$, $(1, 0)^\top$, $(0, 1)^\top$ to the given physical triangle T_ℓ. Recall that, by the chain rule $(\nabla u_h) \circ \Phi_\ell = \boldsymbol{J}_\ell^{-\top} \nabla \hat{u}$ with $\hat{u} = u_h \circ \Phi_\ell$ and thus $\nabla \hat{u}^\top \boldsymbol{J}_\ell^{-1} = ((\nabla u_h) \circ \Phi_\ell)^\top$. Now we have

$$\boldsymbol{u}^\top \boldsymbol{B}_\ell = [\boldsymbol{u}_{\ell,1}, \boldsymbol{u}_{\ell,2}, \boldsymbol{u}_{\ell,3}]\boldsymbol{D}_\ell = \sqrt{\frac{1}{2}\det \boldsymbol{J}_\ell}[\boldsymbol{u}_{\ell,2} - \boldsymbol{u}_{\ell,1}, \boldsymbol{u}_{\ell,3} - \boldsymbol{u}_{\ell,1}]\boldsymbol{J}_\ell^{-1},$$

and noting that $[\boldsymbol{u}_{\ell,2} - \boldsymbol{u}_{\ell,1}, \boldsymbol{u}_{\ell,3} - \boldsymbol{u}_{\ell,1}] = \nabla \hat{u}^\top$ and $\det \boldsymbol{J}_\ell = 2|T_\ell|$ finishes the proof. □

Chosen meshes.
Given a refinement level $n_{\text{ref}} \in \{4, 5, 6\}$, we use a structured mesh with $2^{n_{\text{ref}}} + 1$ many uniformly distributed points per dimension. For $d = 1$ this corresponds to a uniform grid with $n = 2^{n_{\text{ref}}} + 1$ points and $m = 2^{n_{\text{ref}}}$ elements. For $d = 2$, we have $n = (2^{n_{\text{ref}}} + 1)^2$ and $m = 2^{2n_{\text{ref}}+1}$ triangular elements. The triangles are obtained by dividing each square in the Cartesian grid created by the vertices into two triangles with a diagonal connecting the bottom left and top right vertex of a square, see Fig. 1. As a result, our triangular mesh contains only two types of triangles (both being isosceles right triangles): Element type 1 having the right angle on the bottom right, and element type 2 having the right angle on the top left, see Fig. 1.

2.2 Topological Derivative

The topological derivative of a shape function $\mathcal{J} = \mathcal{J}(\Omega)$ represents the sensitivity of $\mathcal{J}$ with respect to a local topological perturbation of the domain Ω around an inner

point z. Consider the setting introduced in Sect. 2.1.1 where Ω denotes a subdomain of a domain D. Let $\omega \subset \mathbb{R}^d$ with $\mathbf{0} \in \omega$ represent the shape of the considered perturbation, e.g., $\omega = B_1(\mathbf{0})$ the unit ball for circular inclusion shapes, and let $z \in \Omega \cup (\mathsf{D} \setminus \overline{\Omega})$. For $\varepsilon > 0$, we define the perturbation of shape ω and size ε as $\omega_\varepsilon(z) := z + \varepsilon\omega$.

Definition 1 (**topological derivative**) The topological derivative of a shape function $\mathcal{J}$ at the point $z \in \Omega \cup (\mathsf{D} \setminus \overline{\Omega})$ with respect to the inclusion shape ω is defined by

$$d\mathcal{J}(\Omega)(z,\omega) := \begin{cases} \lim\limits_{\varepsilon \searrow 0} \frac{1}{|\omega_\varepsilon|}(\mathcal{J}(\Omega \setminus \overline{\omega}_\varepsilon) - \mathcal{J}(\Omega)), & z \in \Omega, \\ \lim\limits_{\varepsilon \searrow 0} \frac{1}{|\omega_\varepsilon|}(\mathcal{J}(\Omega \cup \omega_\varepsilon) - \mathcal{J}(\Omega)), & z \in \mathsf{D} \setminus \overline{\Omega}. \end{cases} \tag{14}$$

Remark 3 Note that Definition 1 is equivalent to stating that $d\mathcal{J}(\Omega)(z,\omega)$ is the first term in a topological asymptotic expansion of the form (here for $z \in \mathsf{D}\setminus\overline{\Omega}$)

$$\mathcal{J}(\Omega \cup \omega_\varepsilon) = \mathcal{J}(\Omega) + |\omega_\varepsilon| d\mathcal{J}(\Omega)(z,\omega) + o(|\omega_\varepsilon|). \tag{15}$$

In general, the topological derivative of PDE-constrained topology optimization problems with elliptic PDE constraints where the principal part of the PDE operator is perturbed involves the solution to an exterior corrector equation, which we define in the following. We refer the reader to [20] for a comprehensive introduction to the numerical computation of topological derivatives for arbitrary inclusion shapes.

Definition 2 Let $\omega \subset \mathbb{R}^d$ open with $0 \in \omega$ and let

$$\lambda_\omega^{i\to j}(x) = \chi_\omega(x)\lambda^j + \chi_{\mathbb{R}^d\setminus\overline{\omega}}(x)\lambda^i \tag{16}$$

for $i, j \in \{\text{in}, \text{out}\}$, $i \neq j$. Furthermore, for $z \in \Omega \cup (\mathsf{D} \setminus \overline{\Omega})$, let $\nabla u(z)$ denote the point evaluation of the gradient of the solution u to (5b) at z. For any given $\zeta \in \mathbb{R}^d$, we define the corrector function $K_\omega[\lambda^i, \lambda^j; \zeta] \in \dot{BL}(\mathbb{R}^d)$ for switching from material λ^i to λ^j, $i, j \in \{\text{in}, \text{out}\}$, $i \neq j$, as the unique solution to

$$\int_{\mathbb{R}^d} \lambda_\omega^{i\to j} \nabla K_\omega[\lambda^i, \lambda^j; \zeta](x) \cdot \nabla v(x)\, \mathrm{d}x = -(\lambda^j - \lambda^i) \int_\omega \zeta \cdot \nabla v(x)\, \mathrm{d}x \tag{17}$$

for all $v \in BL(\mathbb{R}^d)$.

Here, $BL(\mathbb{R}^d) := \{v \in H^1_{\text{loc}}(\mathbb{R}^d) : \nabla v \in L^2(\mathbb{R}^d)^d\}$ denotes the so-called Beppo-Levi space of locally square integrable functions whose gradient is square integrable over the whole unbounded domain and $\dot{BL}(\mathbb{R}^d) := BL(\mathbb{R}^d)/\mathbb{R}$ is the space of equivalence classes where the constants are factored out, see also [21, 22].

Remark 4 Note that, for $i, j \in \{\text{in}, \text{out}\}$, $i \neq j$, the mapping $\zeta \mapsto K_\omega[\lambda^i, \lambda^j; \zeta]$ is linear and we have for $d = 2$

$$K_\omega[\lambda^i, \lambda^j; \nabla u(z)] = \frac{\partial u}{\partial x_1}(z) K_\omega[\lambda^i, \lambda^j; \boldsymbol{e}^{(1)}] + \frac{\partial u}{\partial x_2}(z) K_\omega[\lambda^i, \lambda^j; \boldsymbol{e}^{(2)}].$$

Proposition 2 *Let $\omega \in \mathbb{R}^d$ open with $0 \in \omega$. Let further p be the adjoint state defined in (6) and $K_\omega[\lambda^{\text{in}}, \lambda^{\text{out}}; \cdot]$, $K_\omega[\lambda^{\text{out}}, \lambda^{\text{in}}; \cdot]$ according to Definition 2. The topological derivative of problem (5) with respect to ω for $z \in \mathsf{D}\backslash\overline{\Omega}$ reads*

$$d\mathcal{J}[\lambda^{\text{out}}, \lambda^{\text{in}}](\Omega)(z, \omega) = (\lambda^{\text{in}} - \lambda^{\text{out}})\frac{1}{|\omega|}\int_\omega (\nabla u(z) + \nabla K_\omega[\lambda^{\text{out}}, \lambda^{\text{in}}; \nabla u(z)](x)) \cdot \nabla p(z)\, dx. \quad (18)$$

Likewise, for $z \in \Omega$, the topological derivative is given by

$$d\mathcal{J}[\lambda^{\text{in}}, \lambda^{\text{out}}](\Omega)(z, \omega) = (\lambda^{\text{out}} - \lambda^{\text{in}})\frac{1}{|\omega|}\int_\omega (\nabla u(z) + \nabla K_\omega[\lambda^{\text{in}}, \lambda^{\text{out}}; \nabla u(z)](x)) \cdot \nabla p(z)\, dx. \quad (19)$$

Proof For a detailed proof, see, e.g., [22]. Moreover, the idea of the proof is outlined in Sect. 4.1 where the focus is on triangular inclusion shapes. □

Definition 3 (*weak polarization matrix*) For $\omega \in \mathbb{R}^d$ open with $0 \in \omega$, $\zeta \in \mathbb{R}^d$ and $i, j \in \{\text{in}, \text{out}\}$, $i \neq j$, let $K_\omega[\lambda^i, \lambda^j; \zeta]$ be as defined in Definition 2. We introduce the weak polarization matrix

$$\mathcal{P}_\omega[\lambda^i, \lambda^j] = \left[\frac{1}{|\omega|}\int_\omega \nabla K_\omega[\lambda^i, \lambda^j; \boldsymbol{e}^{(1)}]\mathrm{d}x \quad \frac{1}{|\omega|}\int_\omega \nabla K_\omega[\lambda^i, \lambda^j; \boldsymbol{e}^{(2)}]\mathrm{d}x\right] \in \mathbb{R}^{d\times d}. \quad (20)$$

Using (20) and Remark 4, in the case $z \in \mathsf{D} \setminus \overline{\Omega}$, we can also write (18) as

$$d\mathcal{J}[\lambda^{\text{out}}, \lambda^{\text{in}}](\Omega)(z, \omega) = (\lambda^{\text{in}} - \lambda^{\text{out}})\nabla u(z)^\top(\boldsymbol{I}_2 + \mathcal{P}_\omega[\lambda^{\text{out}}, \lambda^{\text{in}}])\nabla p(z), \quad (21)$$

and an analogous formula is obtained for the case $z \in \Omega$. For a detailed introduction to the concept of polarization tensors, we refer the reader to the book [23].

It can be seen that the evaluation of the topological derivative at a single point z involves the solution of problem (17) with $\zeta = \nabla u(z)$. In the special cases where ω is an interval ($d = 1$), a disk or ellipse ($d = 2$) or a ball or ellipsoid ($d = 3$), the solution to problem (17) can be written explicitly in a closed form. For $d \in \{2, 3\}$ and $\omega = B_1(0)$ we have $\nabla K_\omega[\lambda^i, \lambda^j; \nabla u(z)]|_\omega = -(\lambda^j - \lambda^i)/(\lambda^j + (d-1)\lambda^i)\nabla u(z)$ and thus $\mathcal{P}_\omega[\lambda^i, \lambda^j] = -(\lambda^j - \lambda^i)/(\lambda^j + (d-1)\lambda^i)\boldsymbol{I}_d$ and [24]

$$d\mathcal{J}[\lambda^i, \lambda^j](\Omega)(z, B_1(0)) = d\lambda^i \frac{\lambda^j - \lambda^i}{\lambda^j + (d-1)\lambda^i}\nabla u(z) \cdot \nabla p(z). \quad (22)$$

For $d = 1$ and $\omega = (-1, 1)$ we have $\mathcal{P}_\omega[\lambda^i, \lambda^j] = -(\lambda^j - \lambda^i)/\lambda^j \in \mathbb{R}$ and thus [25]

$$d\mathcal{J}[\lambda^i, \lambda^j](\Omega)(z, B_1(0)) = \frac{\lambda^i}{\lambda^j}(\lambda^j - \lambda^i)u'(z)p'(z).$$

Remark 5 For a general inclusion shape ω, (17) cannot be solved analytically. The same holds true for the case of quasilinear PDE constraints where the problem corresponding to (17) is quasilinear as well. In these cases, however, it is still possible to get a good approximation to the topological derivative values (18), (19) by computing a numerical approximation of the solution of (17) on a comparably large, but bounded domain $B_R(0) \supset \omega$ (e.g. $R = 30$) with homogeneous Dirichlet boundary conditions at $\partial B_R(0)$. This procedure is motivated by the fact that the solution $K_\omega^{i\to j}$ to (17) often can be shown to decay as $|x| \to \infty$. We refer the reader to [20] for a detailed discussion of this aspect. We will also follow this approach in Sect. 4 for triangular shaped inclusions ω.

Note that, when no closed form solution is available, in the case of linear PDE constraints with only two different materials λ^{in}, λ^{out} it is sufficient to have access to (an approximation of) $K_\omega[\lambda^{\text{in}}, \lambda^{\text{out}}; \boldsymbol{e}^{(k)}]$ and $K_\omega[\lambda^{\text{out}}, \lambda^{\text{in}}; \boldsymbol{e}^{(k)}]$ for $k = 1, \ldots, d$ in order to (approximately) evaluate the topological derivative in the full domain in an efficient way. Thus, problem (17) has to be solved numerically only $2d$ many times in a pre-computation stage. For the case of quasilinear PDE constraints this precomputation stage is more involved, see [26, 27].

2.3 Separable Approximations

Problem (2) can be solved efficiently by the idea of sequential global programming (SGP) [12] where the original optimization problem is replaced by a sequence of simpler sub-problems. For these sub-problems it is beneficial to have approximations of the original objective functions, which are separable, since then solving the sub-problem reduces to the solution of several univariate optimization problems. The following definitions can also be found in [19].

Definition 4 (*separable function*) Let $N \in \mathbb{N}$. A function $g : \mathbb{R}^N \to \mathbb{R}$ is called *separable* if there exist functions $g_1, \ldots, g_N : \mathbb{R} \to \mathbb{R}$ and a constant $g_0 \in \mathbb{R}$ such that for all $\boldsymbol{x} \in \mathbb{R}^N$

$$g(\boldsymbol{x}) = g_0 + \sum_{i=1}^{N} g_i(\boldsymbol{x}_i).$$

We call a model g exact if it still coincides with the original function f when only one component is perturbed.

Definition 5 (*separable exact model*) Let $N \in \mathbb{N}, \overline{\boldsymbol{x}} \in \mathbb{R}^N$ and $f : \mathbb{R}^N \to \mathbb{R}$ be given. A separable function $g : \mathbb{R}^N \to \mathbb{R}$ is called a *separable exact model* of f at $\overline{\boldsymbol{x}}$ if

$$g(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)}) = f(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)}) \tag{23}$$

for all $i \in \{1, \ldots N\}$ and all $\delta x \in \mathbb{R}$.

Definition 6 (*separable first order approximation*) Let $N \in \mathbb{N}, \mathcal{I} \subset \mathbb{R}^N, \overline{\boldsymbol{x}} \in \mathcal{I}$ and $f \in C^1(\mathcal{I}, \mathbb{R})$ be given. A function $g \in C^1(\mathcal{I}, \mathbb{R})$ is called a *separable first order approximation* of f at $\overline{\boldsymbol{x}}$ if g is separable and

$$f(\overline{\boldsymbol{x}}) = g(\overline{\boldsymbol{x}}) \qquad \text{and} \qquad \nabla f(\overline{\boldsymbol{x}}) = \nabla g(\overline{\boldsymbol{x}}). \tag{24}$$

Note that if a function g is a separable exact model of a function f it is also a separable first order approximation. In the following lemma we show how, for any given function f, a separable exact model can be constructed.

Lemma 3 *Let $N \in \mathbb{N}$, $\overline{\boldsymbol{x}} \in \mathbb{R}^N$ and $f : \mathbb{R}^N \to \mathbb{R}$ be given. For $j \in \{1, \dots, N\}$ define*

$$g_j(\overline{\boldsymbol{x}}; \cdot) : \mathbb{R} \to \mathbb{R}, \qquad g_j(\overline{\boldsymbol{x}}; s) = f(\overline{\boldsymbol{x}} + (s - \overline{\boldsymbol{x}}_j)\boldsymbol{e}^{(j)}) - f(\overline{\boldsymbol{x}}). \tag{25}$$

Then, g defined by

$$g : \mathbb{R}^N \to \mathbb{R}, \qquad g(\boldsymbol{x}) = f(\overline{\boldsymbol{x}}) + \sum_{j=1}^{N} g_j(\overline{\boldsymbol{x}}; \boldsymbol{x}_j) \tag{26}$$

is a separable exact model of f at $\overline{\boldsymbol{x}}$.

Proof It is obvious from (26) that g is separable. In order to see (23), let $i \in \{1, \dots, N\}$ be fixed. Noting that

$$(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)})_j = \begin{cases} \overline{\boldsymbol{x}}_i + \delta x, & j = i, \\ \overline{\boldsymbol{x}}_j, & j \in \{1, \dots, N\} \setminus \{i\}, \end{cases}$$

we have

$$\begin{aligned} g(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)}) &= f(\overline{\boldsymbol{x}}) + \sum_{j=1}^{N} g_j(\overline{\boldsymbol{x}}; (\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)})_j) \\ &= f(\overline{\boldsymbol{x}}) + g_i(\overline{\boldsymbol{x}}; \overline{\boldsymbol{x}}_i + \delta x) + \sum_{j \neq i} g_j(\overline{\boldsymbol{x}}; \overline{\boldsymbol{x}}_j) \\ &= f(\overline{\boldsymbol{x}}) + f(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)}) - f(\overline{\boldsymbol{x}}) = f(\overline{\boldsymbol{x}} + \delta x\, \boldsymbol{e}^{(i)}), \end{aligned}$$

where we used $g_j(\overline{\boldsymbol{x}}; \overline{\boldsymbol{x}}_j) = 0$. This finishes the proof. □

While Lemma 3 defines a separable exact model g for any function f, it should be noted that the evaluation of g involves $N + 1$ function evaluations of f, which can in practice be prohibitively expensive (in particular when a function evaluation involves the solution of a PDE). Thus, considering the problem at hand (2), our goal in this paper is to find close approximations of the separable exact model defined by (26) which are cheap to evaluate.

2.4 The Sherman–Morrison–Woodbury Formula in a Finite Element Context

The following linear algebra result will prove useful for defining exact models in the context of discretized PDE-constrained material optimization. It gives a formula for the inverse of a perturbed matrix only in terms of the inverse of the unperturbed matrix.

Lemma 4 (Sherman–Morrison–Woodbury formula [28]) *Let $N, k \in \mathbb{N}$ and matrices $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ invertible, $\boldsymbol{U} \in \mathbb{R}^{N\times k}$, $\boldsymbol{V} \in \mathbb{R}^{k\times N}$ be given. It holds*

$$(\boldsymbol{A}+\boldsymbol{U}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I}_k + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U})^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}. \tag{27}$$

Lemma 4 gives rise to a separable exact model of the discrete compliance functional $\mathcal{J} = \mathcal{J}(\boldsymbol{\lambda})$ defined in (12). For that purpose, let the material distribution $\boldsymbol{\lambda} \in [\underline{\lambda}, \overline{\lambda}]^m \in \mathbb{R}^m$ be given and fix an element T_ℓ, $\ell \in \{1, \dots, m\}$. We consider a perturbation of $\boldsymbol{\lambda}$ in this element and define the perturbed vector $\boldsymbol{\eta} := \boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}$ for some $\eta \in [\underline{\lambda}, \overline{\lambda}]$. Note that $\boldsymbol{\eta}$ coincides with $\boldsymbol{\lambda}$ in all components except for component ℓ where it has value η rather than $\boldsymbol{\lambda}_\ell$. Recall the definition of the matrices $\boldsymbol{B}_\ell$, $\ell \in \{1, \dots, m\}$ from Sect. 2.1.2.

Proposition 5 *Let $\boldsymbol{\lambda} \in \mathbb{R}^m$, $\ell \in \{1, \dots, m\}$ fixed and $\boldsymbol{\eta} := \boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}$. Then*

$$\mathcal{J}(\boldsymbol{\eta}) = \mathcal{J}(\boldsymbol{\lambda}) - |T_\ell|(\eta - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top(\boldsymbol{I}_d + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\nabla u_h|_{T_\ell}. \tag{28}$$

Proof The stiffness matrices according to $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ read

$$\boldsymbol{K}(\boldsymbol{\lambda}) = \sum_{k=1}^m \boldsymbol{\lambda}_k B_k B_k^\top \quad \text{and} \quad \boldsymbol{K}(\boldsymbol{\eta}) = \sum_{k=1}^m \boldsymbol{\eta}_k B_k B_k^\top = \boldsymbol{K}(\boldsymbol{\lambda}) + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell \boldsymbol{B}_\ell{}^\top. \tag{29}$$

Thus, the inverse of $\boldsymbol{K}(\boldsymbol{\eta})$ can be obtained by means of Lemma 4 by setting $\boldsymbol{A} = \boldsymbol{K}(\boldsymbol{\lambda}) \in \mathbb{R}^{n\times n}$, $\boldsymbol{U} = (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell \in \mathbb{R}^{n\times d}$ and $\boldsymbol{V} = \boldsymbol{B}_\ell{}^\top \in \mathbb{R}^{d\times n}$ as

$$\boldsymbol{K}(\boldsymbol{\eta})^{-1} = \boldsymbol{K}(\boldsymbol{\lambda})^{-1} - (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell(\boldsymbol{I}_d + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1}. \tag{30}$$

Thus, defining $\tilde{\boldsymbol{u}} = \boldsymbol{K}(\boldsymbol{\eta})^{-1}\boldsymbol{f}$ and using that $\boldsymbol{u} = \boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{f}$ and $\boldsymbol{K}(\boldsymbol{\lambda}) = \boldsymbol{K}(\boldsymbol{\lambda})^\top$, it holds

$$\boldsymbol{f}^\top \tilde{\boldsymbol{u}} = \boldsymbol{f}^\top \boldsymbol{u} - (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{u}^\top \boldsymbol{B}_\ell(\boldsymbol{I}_d + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\boldsymbol{B}_\ell{}^\top \boldsymbol{u}. \tag{31}$$

Noting that $\mathcal{J}(\boldsymbol{\eta}) = \boldsymbol{f}^\top \tilde{\boldsymbol{u}}$, $\mathcal{J}(\boldsymbol{\lambda}) = \boldsymbol{f}^\top \boldsymbol{u}$ and, by Lemma 1, $\boldsymbol{B}_\ell{}^\top \boldsymbol{u} = \sqrt{|T_\ell|}\nabla u_h|_{T_\ell}$ finishes the proof. □

Remark 6 We remark that this procedure can also be followed for non-selfadjoint problems. Using the adjoint state $p_h \in V_h$ whose basis vector $\boldsymbol{p}$ is obtained as the solution of $\boldsymbol{K}(\boldsymbol{\lambda})^\top \boldsymbol{p} = -J'(\boldsymbol{u})$, in the case of a linear cost function J we obtain

$$\begin{aligned}\mathcal{J}(\boldsymbol{\eta}) &= J(\tilde{\boldsymbol{u}}) = J(\boldsymbol{u}) + J'(\boldsymbol{u})(\tilde{\boldsymbol{u}} - \boldsymbol{u}) \\ &= J(\boldsymbol{u}) + J'(\boldsymbol{u})(\boldsymbol{K}(\boldsymbol{\eta})^{-1} - \boldsymbol{K}(\boldsymbol{\lambda})^{-1})\boldsymbol{f}\end{aligned}$$

$$= J(\boldsymbol{u}) - J'(\boldsymbol{u})(\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell(\boldsymbol{I}_d + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\boldsymbol{B}_\ell{}^\top\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{f}$$
$$= \mathcal{J}(\boldsymbol{\lambda}) + |T_\ell|(\eta - \boldsymbol{\lambda}_\ell)(\nabla p_h|_{T_\ell})^\top(\boldsymbol{I}_d + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\nabla u_h|_{T_\ell}$$

If J is not linear the above identity only holds up to a remainder of second order.

From Lemma 3, we get the following result:

Proposition 6 *Let $\boldsymbol{\lambda}, \boldsymbol{\eta} \in \mathbb{R}^m$. The function $\hat{\mathcal{J}}_{SMW}$ defined by*

$$\hat{\mathcal{J}}_{SMW}(\boldsymbol{\eta}) := \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top (\boldsymbol{I}_d + (\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)\boldsymbol{B}_\ell{}^\top\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell)^{-1}\nabla u_h|_{T_\ell} \tag{32}$$

is a separable exact model of $\mathcal{J}$ at $\boldsymbol{\lambda}$.

Proof From Lemma 3 we know that

$$\hat{\mathcal{J}}_{\mathrm{SMW}}(\boldsymbol{\eta}) = \mathcal{J}(\boldsymbol{\lambda}) + \sum_{\ell=1}^{m} \left(\mathcal{J}(\boldsymbol{\lambda} + (\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}) - \mathcal{J}(\boldsymbol{\lambda})\right) \tag{33}$$

is a separable exact model of $\mathcal{J}$ at $\boldsymbol{\lambda}$. Plugging in the result of Proposition 5 yields the assertion. □

Compared to the general separable exact model according to Lemma 3, which can be defined for any function, using the Sherman–Morrison–Woodbury formula we have found a closed form for a separable exact model for our given cost function in (32). Of course, it can be seen that model (32) still involves the inverse of the stiffness matrix for the material distribution given by $\boldsymbol{\lambda}$ which one typically does not have access to. Even when the stiffness matrix has been factorized for computing the state $\boldsymbol{u}$, the evaluation of (32) involves m many forward/backward substitutions which amounts to a total effort in the order of $\mathcal{O}(mn^2)$ and is thus prohibitive for many real-world applications. In the subsequent section, we will introduce an approximation of (32) which can be evaluated efficiently.

3 An Efficient Separable Model Based on the Sherman–Morrison–Woodbury Formula

Recall the separable exact first order model (32) which reads

$$\hat{\mathcal{J}}_{\mathrm{SMW}}(\boldsymbol{\eta}) = \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top \left(\boldsymbol{I}_d - (\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)\boldsymbol{\Gamma}^{(\ell)}\right)^{-1} \nabla u_h|_{T_\ell} \tag{34}$$

with the definition $\boldsymbol{\Gamma}^{(\ell)} := -\boldsymbol{B}_\ell{}^\top\boldsymbol{K}(\boldsymbol{\lambda})^{-1}\boldsymbol{B}_\ell \in \mathbb{R}^{d\times d}$ for all $\ell \in \{1, \ldots, m\}$. Since the evaluation of the model involves the solution of a linear system with the system

matrix $\boldsymbol{K}(\boldsymbol{\lambda})$ for every element index ℓ, we introduce an approximation which can be evaluated more efficiently. For that purpose, we simply approximate the inverse of the stiffness matrix by the inverse of the diagonal approximation of the stiffness matrix to obtain the model

$$\hat{\mathcal{J}}_{\mathrm{SMWdiag}}(\boldsymbol{\eta}) := \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top \left(\boldsymbol{I}_d - (\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)\boldsymbol{\Gamma}_{\mathrm{diag}}^{(\ell)}\right)^{-1} \nabla u_h|_{T_\ell} \quad (35)$$

with

$$\boldsymbol{\Gamma}_{\mathrm{diag}}^{(\ell)} := -\boldsymbol{B}_\ell{}^\top (\mathrm{diag}\boldsymbol{K}(\boldsymbol{\lambda}))^{-1} \boldsymbol{B}_\ell \in \mathbb{R}^{d\times d}. \quad (36)$$

This model is a separable first order model, but it is no longer exact. Note that this idea was already proposed in [19] in the context of a discrete dipole approximation method in an application from optics and is transferred to a finite element setting here.

In the following, we investigate model (35) in the one-dimensional case, where connections to the mathematical concepts of topological and shape derivatives can be established. In spatial dimension $d = 1$, we can compute the matrix $\boldsymbol{\Gamma}_{\mathrm{diag}}^{(\ell)}$ explicitly.

Lemma 7 *Let $d = 1$ and let a uniform mesh $\{T_1, \ldots, T_m\}$ of the computational domain be given. Assume that element T_ℓ is occupied by material λ^{out} and also its two neighbors $T_{\ell-1}$, $T_{\ell+1}$ are occupied by the same material, i.e., $\boldsymbol{\lambda}_{\ell-1} = \boldsymbol{\lambda}_\ell = \boldsymbol{\lambda}_{\ell+1} = \lambda^{\mathrm{out}}$. Then it holds*

$$\boldsymbol{\Gamma}_{diag}^{(\ell)} = -\frac{1}{\lambda^{\mathrm{out}}} \quad (37)$$

and, for $\boldsymbol{\eta} = \boldsymbol{\lambda} + (\lambda^{\mathrm{in}} - \lambda^{\mathrm{out}})\boldsymbol{e}^{(\ell)}$,

$$\hat{\mathcal{J}}_{SMWdiag}(\boldsymbol{\eta}) = \mathcal{J}(\boldsymbol{\lambda}) - |T_\ell| \frac{\lambda^{\mathrm{out}}}{\lambda^{\mathrm{in}}} (\lambda^{\mathrm{in}} - \lambda^{\mathrm{out}})(u_h'|_{T_\ell})^2. \quad (38)$$

Proof Using the definition of $\boldsymbol{B}_\ell$ from Sect. 2.1.2, we have

$$\boldsymbol{B}_\ell{}^\top (\mathrm{diag}\boldsymbol{K}(\boldsymbol{\lambda}))^{-1} \boldsymbol{B}_\ell = h^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}^\top (\tilde{\boldsymbol{B}}_\ell)^\top (\mathrm{diag}\boldsymbol{K}(\boldsymbol{\lambda}))^{-1} \tilde{\boldsymbol{B}}_\ell \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

where h denotes the uniform mesh size. Using that $(\mathrm{diag}\boldsymbol{K}(\boldsymbol{\lambda}))_{ii}^{-1} = \frac{1}{a(\varphi_i,\varphi_i)}$ with $a(\varphi_i, \varphi_i) = \int_{\mathsf{D}} \lambda(x)|\nabla\varphi_i|^2 \,\mathrm{d}x$ and

$$(\tilde{\boldsymbol{B}}_\ell)^\top (\mathrm{diag}\boldsymbol{K}(\boldsymbol{\lambda}))^{-1} \tilde{\boldsymbol{B}}_\ell = \begin{pmatrix} \frac{1}{a(\varphi_{\ell,1},\varphi_{\ell,1})} & 0 \\ 0 & \frac{1}{a(\varphi_{\ell,2},\varphi_{\ell,2})} \end{pmatrix} \quad (39)$$

where $\varphi_{\ell,1}$ and $\varphi_{\ell,2}$ are the basis functions corresponding to the left and right node of element T_ℓ, respectively, we get

$$\boldsymbol{\Gamma}_{\text{diag}}^{(\ell)} = -\boldsymbol{B}_\ell{}^\top (\text{diag}\boldsymbol{K}(\boldsymbol{\lambda}))^{-1} \boldsymbol{B}_\ell = -h^{-1} \left(\frac{1}{a(\varphi_{\ell,1}, \varphi_{\ell,1})} + \frac{1}{a(\varphi_{\ell,2}, \varphi_{\ell,2})} \right). \quad (40)$$

The first result follows by noting that $a(\varphi_{\ell,1}, \varphi_{\ell,1}) = a(\varphi_{\ell,2}, \varphi_{\ell,2}) = \frac{2\lambda^{\text{out}}}{h}$. The second identity follows by plugging in and noting that $\boldsymbol{\eta}_\ell = \lambda^{\text{in}}$, $\boldsymbol{\lambda}_\ell = \lambda^{\text{out}}$. □

Remark 7 (Relation to topological derivative in 1D) As pointed out in Sect. 2.2, the topological derivative of our optimization problem (5) at a point $z \in \mathsf{D}\backslash\overline{\Omega}$ (i.e., where $\lambda_\Omega(z) = \lambda^{\text{out}}$) with respect to $\omega = (-1, 1)$ the one-dimensional unit ball reads

$$d\mathcal{J}[\lambda^{\text{out}}, \lambda^{\text{in}}](\Omega)(z, \omega) = -\frac{\lambda^{\text{out}}}{\lambda^{\text{in}}} (\lambda^{\text{in}} - \lambda^{\text{out}})(u'(z))^2, \quad (41)$$

where we used that $p = -u$ for our particular problem at hand. This means that, in one space dimension, the model $\hat{\mathcal{J}}_{\text{SMWdiag}}$ introduced in (35) actually coincides with the finite element discretization of the model that is naturally defined by the definition of the topological derivative

$$\mathcal{J}(\Omega \cup \omega_\varepsilon) \approx \mathcal{J}(\Omega) + |\omega_\varepsilon| d\mathcal{J}(\Omega)(z, \omega), \quad (42)$$

see also (15).

We mention that this direct correspondence of the discrete model $\hat{\mathcal{J}}_{\text{SMWdiag}}$ and the closed-form formula of the topological derivative only works since the elements T_ℓ in a 1D mesh are scaled versions of the 1D unit ball $\omega = B_1(0) = (-1, 1)$. This is no longer the case in two or three dimensions, where elements are polygonal or polyhedral.

We remark that we also observed numerically that the finite element discretization of (42) and the model (35) coincide in elements T_ℓ in homogeneous regions (i.e., where $\boldsymbol{\lambda}_{\ell-1} = \boldsymbol{\lambda}_\ell = \boldsymbol{\lambda}_{\ell+1}$). In elements T_ℓ that are adjacent to the material interface $\partial\Omega$, however, Lemma 7 and thus formula (38) are no longer valid. We observed that the topological derivative model (42), however, still yielded very good results. This can be explained by the following discussion on the relation between the topological and shape derivative in 1D.

Remark 8 (Relation to shape derivative in 1D) The shape derivative for moving the interface $\Gamma := \overline{\Omega} \cap \overline{(\mathsf{D}\backslash\Omega)}$ in the direction given by a vector field $V \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ can, by the structure theorem of Hadamard-Zolesio [1], always (under suitable smoothness assumptions) be written in the form

$$d\mathcal{J}(\Omega; V) = \int_\Gamma L(V \cdot n)\, ds_x,$$

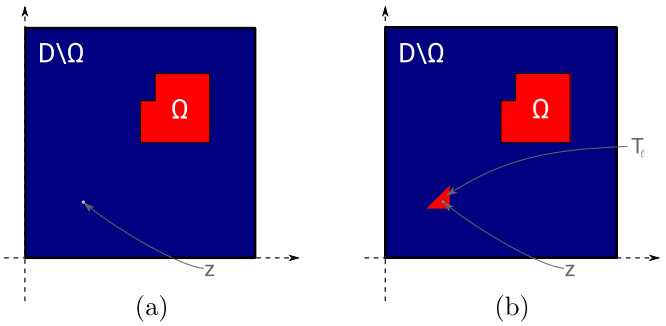


Fig. 2 **a** Unperturbed configuration. **b** Perturbed configuration where the domain is perturbed in triangle T_ℓ whose centroid is the point z

for some scalar function L. Here $\Omega \subset \mathsf{D}$ is the domain where $\lambda = \lambda^{\text{in}}$ and $\mathsf{D}\backslash\overline{\Omega}$ is where $\lambda = \lambda^{\text{out}}$, and n denotes the outer unit normal vector to Ω. For our problem (5), L is given by the formula

$$L = (\lambda^{\text{in}} - \lambda^{\text{out}})(\nabla u \cdot \tau)(\nabla p \cdot \tau) - \left(\frac{1}{\lambda^{\text{in}}} - \frac{1}{\lambda^{\text{out}}}\right)(\lambda_\Omega \nabla u \cdot n)(\lambda_\Omega \nabla p \cdot n)$$

with the tangential vector τ (see, e.g., [29]). In the case where $d = 1$ the tangential derivative of u and thus the first term of L vanishes. Using that n is a scalar with $n^2 = 1$ and again that $p = -u$, we get

$$L = \left(\frac{1}{\lambda^{\text{in}}} - \frac{1}{\lambda^{\text{out}}}\right)(\lambda^{\text{out}} u'_{\text{out}}|_\Gamma)^2 = -\frac{\lambda^{\text{out}}}{\lambda^{\text{in}}}(\lambda^{\text{in}} - \lambda^{\text{out}})(u'_{\text{out}}|_\Gamma)^2,$$

where $u'_{\text{out}}|_\Gamma$ denotes the limit at Γ of the discontinuous quantity u' when coming from $\mathsf{D}\backslash\overline{\Omega}$.

Note that this formula for the 1D shape derivative resembles the topological derivative formula (41), which explains why (42) is a very good model in 1D.

4 A Separable Model Based on the Topological Derivative

In this section, we propose a separable model that is based on the notion of the topological derivative. We fix the space dimension $d = 2$. The topological derivative of a shape function $\mathcal{J} = \mathcal{J}(\Omega)$ with respect to a perturbation of shape ω around a spatial point z was introduced in Sect. 2.2. We emphasize that, while closed-form formulas for the topological derivative only exist in the case of circular or elliptic inclusion shapes [24, 30, 31], a numerical approximation of the weak polarization matrix (20) and thus of the topological derivative formulas (18) and (19) is possible for arbitrary inclusion shapes ω with $0 \in \omega$, see also [20]. We will follow this idea for the case of triangular inclusion shapes.

Let $\Omega \subset \mathsf{D}$ be given and consider an element $T_\ell \in \mathcal{T}$ of type 1 (cf. Figure 1) with $T_\ell \subset \mathsf{D} \setminus \overline{\Omega}$ with vertices $\boldsymbol{x}_{\ell,1}, \boldsymbol{x}_{\ell,2}, \boldsymbol{x}_{\ell,3}$ (in counter-clockwise enumeration) and centroid $\boldsymbol{z}_\ell := (\boldsymbol{x}_{\ell,1} + \boldsymbol{x}_{\ell,2} + \boldsymbol{x}_{\ell,3})/3$, see Fig. 2. Let $\hat{T}^{(1)}$ denote the reference triangle defined by its three vertices $\frac{1}{3}(-2,-1)^\top$, $\frac{1}{3}(1,-1)^\top$, $\frac{1}{3}(1,2)^\top$ and $\hat{T}^{(2)}$ the reference triangle with vertices $\frac{1}{3}(2,1)^\top$, $\frac{1}{3}(-1,-2)^\top$, $\frac{1}{3}(-1,1)^\top$. For the rest of this section we restrict ourselves to elements of type 1. We set $\hat{T} := \hat{T}^{(1)}$ and define $\Phi_{h,\ell} : \hat{T} \to T_\ell, \boldsymbol{x} \mapsto \boldsymbol{z}_\ell + h\boldsymbol{x}$ as the affine transformation satisfying $\Phi_{h,\ell}(\hat{T}) = T_\ell$. We remark that the procedure is completely analogous for elements of type 2 using reference triangle $\hat{T}^{(2)}$.

4.1 Derivation of Topological Derivatives for Triangular Inclusion Shapes

Given a domain Ω, recall the notation $\lambda_\Omega(x) = \chi_\Omega(x)\lambda^{\text{in}} + \chi_{\mathsf{D}\setminus\overline{\Omega}}(x)\lambda^{\text{out}}$. For the fixed triangular domain perturbation T_ℓ, let the perturbed solution $u^{(\ell)} \in V_g$ be defined as the unique solution satisfying

$$\int_{\mathsf{D}} \lambda_{\Omega\cup T_\ell}(x)\nabla u^{(\ell)} \cdot \nabla v \,\mathrm{d}x = \int_{\mathsf{D}} f v \,\mathrm{d}x + \int_{\Gamma_N} g_N v \,\mathrm{d}s_x \tag{43}$$

for all $v \in V_0$. We rewrite the difference of the perturbed and unperturbed cost function by adding the equations (43) and (5b) defining $u^{(\ell)}$ and u, respectively. Using the adjoint state p defined by (6) as test function, we obtain

$$\begin{aligned}
\mathcal{J}(\Omega \cup T_\ell) - \mathcal{J}(\Omega) &= J(u^{(\ell)}) - J(u) \\
&= J(u^{(\ell)}) + \int_{\mathsf{D}} \lambda_{\Omega\cup T_\ell}(x)\nabla u^{(\ell)} \cdot \nabla p \,\mathrm{d}x - \int_{\mathsf{D}} fp \,\mathrm{d}x - \int_{\Gamma_N} g_N p \,\mathrm{d}s_x \\
&\quad - J(u) - \int_{\mathsf{D}} \lambda_\Omega(x)\nabla u \cdot \nabla p \,\mathrm{d}x + \int_{\mathsf{D}} fp \,\mathrm{d}x + \int_{\Gamma_N} g_N p \,\mathrm{d}s_x \\
&= \int_{\mathsf{D}} f(u^{(\ell)} - u) \,\mathrm{d}x + \int_{\Gamma_N} g_N (u^{(\ell)} - u) \,\mathrm{d}s_x \\
&\quad + \int_{\mathsf{D}} \lambda_\Omega(x)\nabla(u^{(\ell)} - u) \cdot \nabla p \,\mathrm{d}x + (\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{T_\ell} \nabla u^{(\ell)} \cdot \nabla p \,\mathrm{d}x \\
&= (\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{T_\ell} \nabla(u^{(\ell)} - u) \cdot \nabla p \,\mathrm{d}x + (\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{T_\ell} \nabla u \cdot \nabla p \,\mathrm{d}x
\end{aligned}$$

where we used the adjoint equation (6) in the last step. Making a change of variables $x \mapsto \Phi_{h,\ell}(x)$ and defining $K_{h,\ell,\hat{T}}[\lambda^{\text{out}}, \lambda^{\text{in}}] := \frac{1}{h}(u^{(\ell)} - u) \circ \Phi_{h,\ell}$, we get

$$\begin{aligned}
\mathcal{J}(\Omega \cup T_\ell) = \mathcal{J}(\Omega) &+ h^2(\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{\hat{T}} \nabla K_{h,\ell,\hat{T}}[\lambda^{\text{out}}, \lambda^{\text{in}}] \cdot (\nabla p) \circ \Phi_{h,\ell} \,\mathrm{d}x \\
&+ h^2(\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{\hat{T}} (\nabla u) \circ \Phi_{h,\ell} \cdot (\nabla p) \circ \Phi_{h,\ell} \,\mathrm{d}x,
\end{aligned} \tag{44}$$

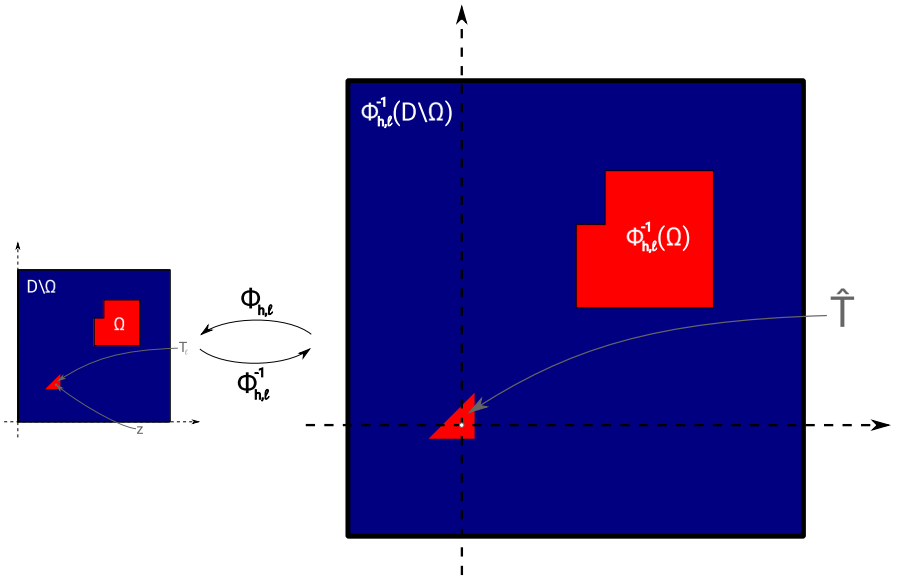


Fig. 3 Rescaled perturbed domain

where we used that, according to the chain rule, $(\nabla v) \circ \Phi_{h,\ell} = \frac{1}{h}\nabla(v \circ \Phi_{h,\ell})$ and $\det(\partial \Phi_{h,\ell}) = h^2$.

Subtracting the unperturbed state equation (5b) from the perturbed equation (43), we see that $u^{(\ell)} - u \in V_0$ satisfies

$$\int_{\mathsf{D}} \lambda_{\Omega \cup T_\ell} \nabla(u^{(\ell)} - u) \cdot \nabla v \, \mathrm{d}x = -(\lambda^{\mathrm{in}} - \lambda^{\mathrm{out}}) \int_{T_\ell} \nabla u \cdot \nabla v \, \mathrm{d}x$$

for all $v \in V_0$. Making the same change of variables, this amounts to

$$\begin{aligned} &\int_{\Phi_{h,\ell}^{-1}(\mathsf{D})} \lambda_{\hat{T} \cup \Phi_{h,\ell}^{-1}(\Omega)} \nabla K_{h,\ell,\hat{T}}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}] \cdot \nabla \psi \, \mathrm{d}x \\ &\quad = -(\lambda^{\mathrm{in}} - \lambda^{\mathrm{out}}) \int_{\hat{T}} (\nabla u) \circ \Phi_{h,\ell} \cdot \nabla \psi \, \mathrm{d}x \end{aligned} \tag{45}$$

for all $\psi \in H_0^1(\Phi_{h,\ell}^{-1}(\mathsf{D}))$. The domain and material distribution of problem (45) is depicted in Fig. 3.

Obviously, if $K_{h,\ell,\hat{T}}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}]$ was known exactly then (44) would give rise to an exact model for $\mathcal{J}(\Omega \cup T_\ell)$. However, of course, this would require the solution of a linear problem for every element ℓ and is therefore computationally not tractable. Instead, we now aim at obtaining an approximation of the quantity $K_{h,\ell,\hat{T}}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}]$ that is independent of the particular element index ℓ. As it is often used in the derivation of topological derivatives, we consider the limit of problem (45) as $h \to 0$. This leads to the problem to find $K_{\hat{T}}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}; \nabla u(z_\ell)] \in X$ satisfying [22]

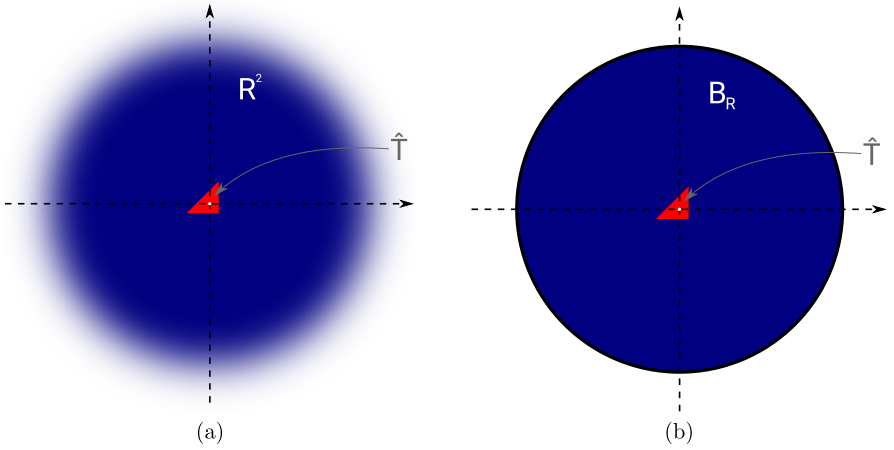


Fig. 4 **a** Rescaled perturbed domain after limit $h \to 0$. **b** Truncation of (**a**) at radius R

$$\int_{\mathbb{R}^2} \lambda_{\hat{T}} \nabla K_{\hat{T}}[\lambda^{\text{out}}, \lambda^{\text{in}}; \nabla u(z_\ell)] \cdot \nabla \psi \, \mathrm{d}x = -(\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{\hat{T}} \nabla u(z_\ell) \cdot \nabla \psi \, \mathrm{d}x \quad (46)$$

for all test functions $\psi \in X$ where X is a Beppo-Levi space, see Sect. 2.2. For an illustration of the corresponding material distribution, see Fig. 4a. Note that, if we had access to the exact solution $K_{\hat{T}}[\lambda^{\text{out}}, \lambda^{\text{in}}; \nabla u(z_\ell)]$ of (46), the topological derivative at the centroid z_ℓ of triangle $T_\ell \in \mathsf{D} \backslash \overline{\Omega}$ with respect to $\hat{T}$-shaped inclusion shapes would follow from (44) as

$$\begin{aligned} & d\mathcal{J}[\lambda^{\text{out}}, \lambda^{\text{in}}](\Omega)(z_\ell, \hat{T}) \\ & \quad = \lim_{h \to 0} \frac{\mathcal{J}(\Omega \cup T_\ell) - \mathcal{J}(\Omega)}{|T_\ell|} \\ & \quad = (\lambda^{\text{in}} - \lambda^{\text{out}}) \frac{1}{|\hat{T}|} \int_{\hat{T}} (\nabla u(z_\ell) + \nabla K_{\hat{T}}[\lambda^{\text{out}}, \lambda^{\text{in}}; \nabla u(z_\ell)](x)) \cdot \nabla p(z_\ell) \, \mathrm{d}x \end{aligned} \quad (47)$$

which coincides with the statement of Proposition 2. Here we used $|T_\ell| = h^2 |\hat{T}|$.

4.2 Our Proposed Topological Derivative Model

Unlike in the case of circular or elliptic inclusions, no closed-form solution to the exterior problem (46) for triangular inclusion shapes $\omega = \hat{T}$ is known in the literature. Thus, formula (47) cannot be evaluated exactly. However, as it was shown in [20], it is feasible to numerically approximate the exterior problem (46) by truncating the domain at a moderately large radius R (e.g., $R = 30$) and using a finite element discretization with homogeneous Dirichlet boundary conditions on the boundary of the truncated domain (see Fig. 4b). We remark that this truncation is justified, since it is known that the solution to (46) exhibits a certain decay as $|x| \to \infty$ [20].

4.2.1 Topological Derivative Model in Homogeneous Regions

We restrict ourselves to elements T_ℓ in the interior of $\mathsf{D}\setminus\overline{\Omega}$ such that also all neighboring elements of T_ℓ (i.e., elements that share at least one vertex with T_ℓ) are in $\mathsf{D}\setminus\overline{\Omega}$. Of course, all results and statements follow analogously for elements T_ℓ in the interior of Ω. For this setting, we propose the model that is based on the following procedure:

1. Compute a finite element approximation of (46) using a finite element discretization of a truncated domain. More precisely, given a truncation radius R and a mesh $\{\tau_1,\dots,\tau_M\}$ of the truncated domain $B_R(0)$ that resolves the inclusion $\hat{T}$, we aim to find $K_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}};\boldsymbol{e}^{(k)}]\in V_h^R:=\{v\in C^0(B_R(0)):v|_{\partial B_R(0)}=0, v|_{\tau_i}\in P^1, i=1,\dots,M\}$ such that

$$\int_{B_R(0)}\lambda_{\hat{T}}\nabla K_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}};\boldsymbol{e}^{(k)}]\cdot\nabla\psi_h\,\mathrm{d}x=-(\lambda^{\text{in}}-\lambda^{\text{out}})\int_{\hat{T}}\boldsymbol{e}^{(k)}\cdot\nabla\psi_h\,\mathrm{d}x \tag{48}$$

   for all $\psi_h\in V_h^R$ for $k=1$ and $k=2$. Here, recall that $\lambda_{\hat{T}}(x)=\chi_{\hat{T}}(x)\lambda^{\text{in}}+\chi_{\mathbb{R}^2\setminus\hat{T}}(x)\lambda^{\text{out}}$.
2. Compute the approximate weak polarization matrix

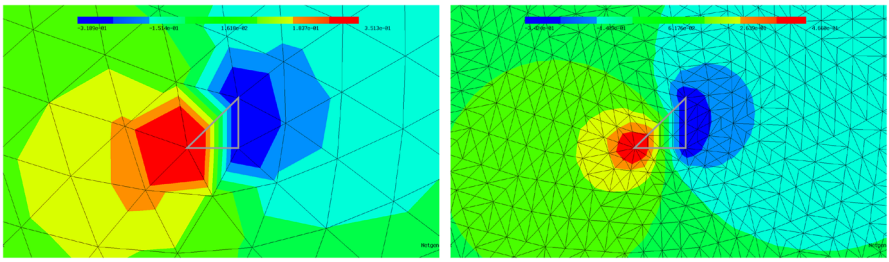
$$\mathcal{P}_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}}]=\left[\frac{1}{|\hat{T}|}\int_{\hat{T}}\nabla K_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}};\boldsymbol{e}^{(1)}]\mathrm{d}x\quad\frac{1}{|\hat{T}|}\int_{\hat{T}}\nabla K_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}};\boldsymbol{e}^{(2)}]\mathrm{d}x\right]\in\mathbb{R}^{2\times 2}. \tag{49}$$

3. Evaluate $d\mathcal{J}_h[\lambda^{\text{out}},\lambda^{\text{in}}](z_\ell,\hat{T})=-(\lambda^{\text{in}}-\lambda^{\text{out}})(\nabla u_h|_{T_\ell})^\top\left(\boldsymbol{I}_2+\mathcal{P}_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}}]\right)\nabla u_h|_{T_\ell}$.

With this, for a given material distribution $\boldsymbol{\lambda}\in\mathbb{R}^m$ and $\boldsymbol{\eta}=\boldsymbol{\lambda}+(\eta-\boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}$ where T_ℓ is in the interior of $\mathsf{D}\setminus\overline{\Omega}$, we obtain the approximation

$$\begin{aligned}\mathcal{J}(\boldsymbol{\eta})\approx&\mathcal{J}(\boldsymbol{\lambda})+|T_\ell|d\mathcal{J}_h[\boldsymbol{\lambda}_\ell,\eta](z_\ell,\hat{T})\\=&\mathcal{J}(\boldsymbol{\lambda})-|T_\ell|(\eta-\boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top\left(\boldsymbol{I}_2+\mathcal{P}_{\hat{T},h}[\boldsymbol{\lambda}_\ell,\eta]\right)\nabla u_h|_{T_\ell}.\end{aligned}$$

Remark 9 Concerning the numerical solution of (48), we make one important remark. It is essential that the mesh $\{\tau_1,\dots,\tau_M\}$ is chosen in such a way that the triangle $\hat{T}$ is discretized by exactly one element τ_j and that, thus, the solution is linear inside the whole of $\hat{T}$. While a finer discretization of $\hat{T}$ would yield a better approximation to the true solution of limit problem (46), the term $K_{\hat{T},h}[\lambda^{\text{out}},\lambda^{\text{in}};\boldsymbol{e}^{(k)}]$ should actually make up for the error $u_h^{(\ell)}-u_h$ inside element T_ℓ which is a linear function inside T_ℓ due to the chosen discretization. Figure 5 shows the solution to (48) when $\hat{T}$ is resolved by exactly one element and when the whole mesh is twice uniformly refined.



(a) (b)

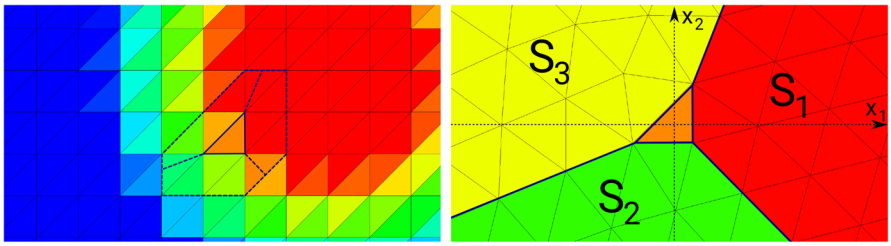
Fig. 5 Comparison of numerical solution $K_{\hat{T},h}[1, 1000; \boldsymbol{e}^{(1)}]$ to (48) on a mesh where $\hat{T}$ is resolved by exactly one element (**a**) and on a twice uniformly refined mesh. The mesh in (a) should be used for solving (48)

4.2.2 Topological Derivative Model in Inhomogeneous Regions

Even if one is interested in binary designs without intermediate materials, in the course of a density-based topology or material optimization procedure, one will of course always encounter regions of intermediate materials. For elements T_ℓ in these regions, the assumptions taken at the beginning of Sect. 4.2.1 are not satisfied and the corresponding proposed model will not be very accurate in these regions. In order to improve the quality of the approximation also in these regions, we recall the idea behind the topological derivative model: The quantity $K_{\hat{T}}$ should approximate the local variation of the (discretized) state with respect to a material perturbation in some element T_ℓ, i.e., it should approximate $u_h^{(\ell)} - u_h$. In other words, problem (46) can be interpreted as considering $u_h^{(\ell)} - u_h$ and zooming in around the fixed triangle T_ℓ and neglecting everything that is beyond a certain (small) distance from that triangle, see also the illustrations in Figs. 3 and 4.

We follow this idea also in the case of inhomogeneous material around a fixed triangle, i.e., we want to approximate the local material distribution in a truncated rescaled domain $B_R(0)$ similar to Fig. 4b. For obtaining an approximation of the inhomogeneous material distribution within the computational domain D, we divide the domain $B_R(0)$ into three sectors. The sectors are separated by three lines which are chosen such as to halve the three interior angles of the triangle $\hat{T}$. Thus, we end up with a domain $B_R(0)$ similar to the one depicted in Fig. 4b which is occupied by four different materials (one inside the triangle $\hat{T}$ and one in each of the three sectors), see Fig. 6b. For the computation of average values within one of the sectors, we take a weighted Hölder average of the values in the neighboring elements with parameter α, i.e., the averaged value in Sector j, $j = 1, 2, 3$, is chosen as

$$\lambda_{T_\ell}^{S_j} = \left(\sum_{T \in \mathcal{N}(T_\ell)} w_{S_j,T} (\boldsymbol{\lambda}_T)^\alpha \right)^{\frac{1}{\alpha}}. \tag{50}$$



(a) (b)

Fig. 6 **a** Element inside inhomogeneous material distribution in computational domain D. Neighborhood for averaging into sector values is marked. **b** Averaged material distribution in three sectors of truncated unbounded domain $B_R(0)$. The average values per sector are obtained by a weighted Hölder mean of the material values in the neighboring elements

Here, $\mathcal{N}(T_\ell)$ denotes the set of triangles that have at least one common vertex with triangle T_ℓ, see Fig. 6a, and $w_{S_j,T} = |T \cap S_j|/|T| \in [0, 1]$ is the volume fraction of triangle T in Sector S_j. Moreover, $\boldsymbol{\lambda}_T$ denotes the entry of the vector $\boldsymbol{\lambda}$ corresponding to the triangle T. In our experiments, we chose the Hölder parameter as $\alpha = -0.5$. This choice will be motivated later in Remark 13 of Sect. 7.

Our proposed model in the case of inhomogeneous material around an element T_ℓ with material coefficient $\boldsymbol{\lambda}_\ell$ and averaged sector values $\lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}$ according to (50) follows the same three steps outlined for the homogeneous setting in Sect. 4.2.1: We numerically compute the corresponding correctors $K_{\hat{T},h}[(\lambda^{\text{out}}, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}), \lambda^{\text{in}}; \boldsymbol{e}^{(k)}]$ for $k = 1, 2$ as the finite element solutions to (48) where $\lambda_{\hat{T}}$ is replaced by the new three-sector material distribution

$$\lambda_{\hat{T}}(x) = \chi_{\hat{T}}(x)\lambda^{\text{in}} + \sum_{j=1}^{3} \chi_{S_j}(x)\lambda_{T_\ell}^{S_j}. \tag{51}$$

Subsequently, the corresponding weak polarization matrix $\mathcal{P}_{\hat{T},h}[(\lambda^{\text{out}}, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}), \lambda^{\text{in}}]$ and the quantity $d\mathcal{J}_h[(\lambda^{\text{out}}, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}), \lambda^{\text{in}}](z_\ell, \hat{T})$ can be computed according to steps 2 and 3 of Sect. 4.2.1.

Summarizing, we define the model

$$\begin{aligned}\hat{\mathcal{J}}_{\text{TDnum}}(\boldsymbol{\eta}) :=& \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top \\ & \left(\boldsymbol{I}_2 + \mathcal{P}_{\hat{T},h}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}), \boldsymbol{\eta}_\ell]\right) \nabla u_h|_{T_\ell}.\end{aligned} \tag{52}$$

Note that the homogeneous setting of Sect. 4.2.1 is actually also covered by the more general inhomogeneous setting treated here.

We next make an important remark concerning the efficient evaluation of the model $\hat{\mathcal{J}}_{\text{TDnum}}$.

Remark 10 In the general inhomogeneous setting, the procedure described in this section states that, in order to evaluate model (52), a problem of type (48) has to be solved for each element T_ℓ. Of course, this is computationally expensive and therefore not recommended by the authors. Instead, the approach followed here is to divide the numerical computations into an offline and an online stage. In the offline stage, which has to be performed only once for the particular type of PDE operator, we compute the quantities

$$K_{\hat{T},h}[(\lambda^{\text{out}}, \lambda^{S_1}, \lambda^{S_2}, \lambda^{S_3}), \lambda^{\text{in}}; \boldsymbol{e}^{(k)}]$$

for $k = 1, 2$ and for a large number of combinations of relevant values $(\lambda^{\text{out}}, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}, \lambda^{\text{in}})$ and store the corresponding weak polarization matrices. This, initially, results in a five-dimensional array of 2×2 matrices. Moreover, these precomputations should be done for each type of reference triangle, i.e., in our case for $\hat{T} = \hat{T}^{(1)}$ and $\hat{T} = \hat{T}^{(2)}$, see Fig. 1.

In the online stage, for each element T_ℓ the corresponding average sector values are computed according to (50) and the weak polarization matrix $\mathcal{P}_{\hat{T},h}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}), \boldsymbol{\eta}_\ell]$ is approximately obtained by piecewise linear interpolation of the precomputed values.

We finally remark that the precomputation can be reduced from five to four dimensions by exploiting that problem (48) with $\lambda_{\hat{T}}$ according to (51) depends on the parameter λ^{out} only via the scaling of the right hand side by $(\lambda^{\text{in}} - \lambda^{\text{out}})$.

5 An Improved Separable Model Based on the Sherman–Morrison–Woodbury Formula

In this section, we revisit the separable exact model defined in (32) and have a closer look at the matrix

$$\boldsymbol{\Gamma}^{(\ell)} = -\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{B}_\ell \in \mathbb{R}^{2\times 2}.$$

Recall that, in order to employ model (32), this matrix would have to be evaluated for each element index ℓ, which amounts to solving m many systems of linear equations and is thus computationally prohibitive. Motivated by the procedure of Sect. 4, our goal here is to find a good approximation of $\boldsymbol{\Gamma}^{(\ell)}$ that is independent of the element index ℓ and can thus be precomputed in an offline stage.

We begin by making the following observation. We assume the finite element setting introduced in Sect. 2.1.2 with the mesh $\mathcal{T}$ and the finite element space $V_h \subset H^1_{\Gamma_D}(\mathsf{D})$ of piecewise linear and globally continuous functions.

Lemma 8 *Let $T_\ell \in \mathcal{T}$ and, for $k = 1, 2$, define $w_{k,h} \in V_h$ the unique numerical solution to the variational problem*

$$\int_D \lambda(x) \nabla w_{k,h} \cdot \nabla v_h \, dx = -\int_{T_\ell} \boldsymbol{e}^{(k)} \cdot \nabla v_h \, dx \tag{53}$$

for all $v_h \in V_h$. Then it holds

$$\mathbf{\Gamma}^{(\ell)} = [\nabla w_{1,h}|_{T_\ell} \ \nabla w_{2,h}|_{T_\ell}]. \tag{54}$$

Proof We use the notation and symbols introduced in Sect. 2.1.2. The discretization of (53) reads

$$\boldsymbol{K}(\boldsymbol{\lambda})\boldsymbol{w}^{(k)} = \boldsymbol{f}^{(\ell,k)}$$

where $\boldsymbol{K}(\boldsymbol{\lambda})$ is the invertible stiffness matrix, $\boldsymbol{w}^{(k)}$ denotes the coefficient vector of the finite element function $w_{k,h} = \sum_{i=1}^{n} \boldsymbol{w}_i^{(k)} \varphi_i$, $k \in \{1, 2\}$, and $\boldsymbol{f}^{(\ell,k)} \in \mathbb{R}^n$ with

$$(\boldsymbol{f}^{(\ell,k)})_i = -\int_{T_\ell} \boldsymbol{e}^{(k)} \cdot \nabla\varphi_i \, \mathrm{d}x, \ i = 1, \ldots n.$$

Since the global load vector $\boldsymbol{f}^{(\ell,k)}$ has contributions only from one element, it holds $\boldsymbol{f}^{(\ell,k)} = \tilde{\boldsymbol{B}}_\ell \boldsymbol{f}_{\mathrm{loc}}^{\ell,k}$ with the element load vector

$$\boldsymbol{f}_{\mathrm{loc}}^{\ell,k} = -|T_\ell| \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{J}_\ell^{-1} \boldsymbol{e}^{(k)} = -\sqrt{|T_\ell|}\boldsymbol{D}_\ell \boldsymbol{e}^{(k)}.$$

Thus, it holds

$$\boldsymbol{B}_\ell{}^\top \boldsymbol{w}^{(k)} = \boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{f}^{(\ell,k)} = \boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \tilde{\boldsymbol{B}}_\ell \boldsymbol{f}_{\mathrm{loc}}^{\ell,k} = -\sqrt{|T_\ell|}\boldsymbol{B}_\ell{}^\top \boldsymbol{K}(\boldsymbol{\lambda})^{-1} \boldsymbol{B}_\ell \boldsymbol{e}^{(k)}. \tag{55}$$

On the other hand, we know from (13) that

$$\boldsymbol{B}_\ell{}^\top \boldsymbol{w}^{(k)} = \sqrt{|T_\ell|}\, \nabla w_{k,h}|_{T_\ell}. \tag{56}$$

Comparing (55) and (56) for $k = 1$ and $k = 2$ yields the result. □

Lemma 8 gives an interpretation of the matrix $\mathbf{\Gamma}^{(\ell)}$, which appears in the Sherman–Morrison–Woodbury model (32) and is costly to evaluate, in terms of a boundary value problem. In order to find an approximation of $\mathbf{\Gamma}^{(\ell)}$ that is independent of the element index ℓ, we proceed similarly to Sect. 4. In boundary value problem (53), we zoom in around the element T_ℓ, i.e., we apply the transformation $\Phi_{h,\ell}^{-1}$ that transforms T_ℓ to the reference element $\hat{T}$, see Fig. 7 for an illustration in a homogeneous setting. Note that, as opposed to the procedure in Sect. 4, here an unperturbed material distribution is transformed.

Passing to the limit $h \to 0$ yields an exterior problem on the unbounded domain, see Fig. 8(a) and again truncating this domain leads to the boundary value problem

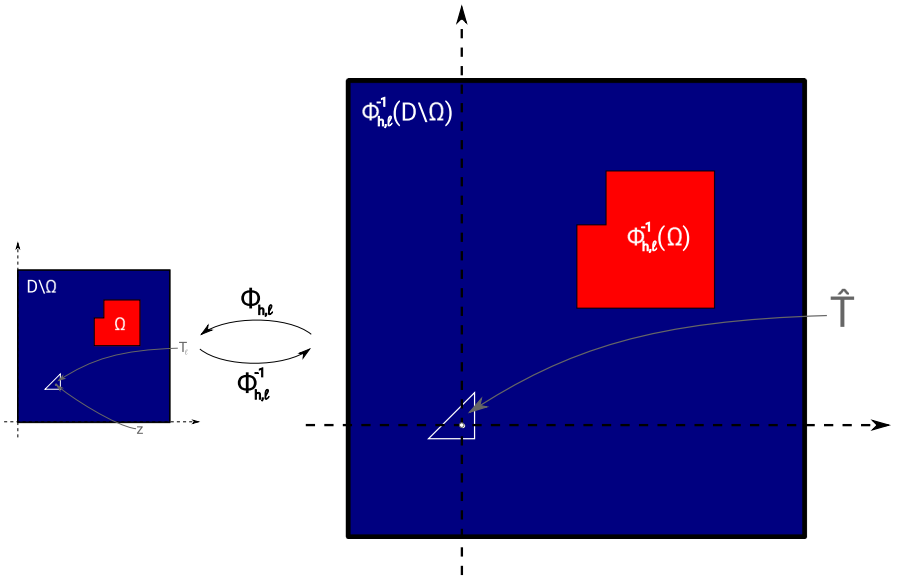


Fig. 7 Rescaled unperturbed domain

on the truncated domain $B_R(0)$ to find $W_{\hat{T}}[\lambda^{\text{out}}; \boldsymbol{e}^{(k)}] \in H_0^1(B_R(0))$, $k = 1, 2$, as the unique solution to

$$\int_{B_R(0)} \lambda^{\text{out}} \nabla W_{\hat{T}}[\lambda^{\text{out}}; \boldsymbol{e}^{(k)}] \cdot \nabla \psi \, \mathrm{d}x = -\int_{\hat{T}} \boldsymbol{e}^{(k)} \cdot \nabla \psi \, \mathrm{d}x \tag{57}$$

for all $\psi \in H_0^1(B_R(0))$, see Fig. 8(b). Note that this problem differs from problem (48) only by a different scaling factor on the right hand side and by a homogeneous material distribution λ^{out}.

In the case when T_ℓ is in an inhomogeneous region of the computational domain D (i.e., not all neighbors of T_ℓ have the same material coefficient), we can follow the same averaging procedure with three sectors as in Sect. 4.2.2 and obtain the problem to find $W_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}); \boldsymbol{e}^{(k)}] \in H_0^1(B_R(0))$, $k = 1, 2$, as the unique solution to

$$\int_{B_R(0)} \lambda_{\hat{T}} \nabla W_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}); \boldsymbol{e}^{(k)}] \cdot \nabla \psi \, \mathrm{d}x = -\int_{\hat{T}} \boldsymbol{e}^{(k)} \cdot \nabla \psi \, \mathrm{d}x \tag{58}$$

where $\lambda_{\hat{T}}(x) = \chi_{\hat{T}}(x)\boldsymbol{\lambda}_\ell + \sum_{j=1}^3 \chi_{S_j}(x)\lambda_{T_\ell}^{S_j}$, cf. also the material distribution in Fig. 6. Remark 9 concerning the numerical approximation of (57) and (58) with a mesh where the subdomain $\hat{T}$ of $B_R(0)$ is discretized by exactly one element remains valid. We define the 2×2 matrix

$$\begin{aligned} \boldsymbol{\Gamma}_{\hat{T},\ell} &:= \boldsymbol{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})] \\ &:= \left[\frac{1}{\hat{T}} \int_{\hat{T}} \nabla W_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}); \boldsymbol{e}^{(1)}] \mathrm{d}x \quad \frac{1}{\hat{T}} \int_{\hat{T}} \nabla W_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}); \boldsymbol{e}^{(2)}] \mathrm{d}x \right], \end{aligned} \tag{59}$$

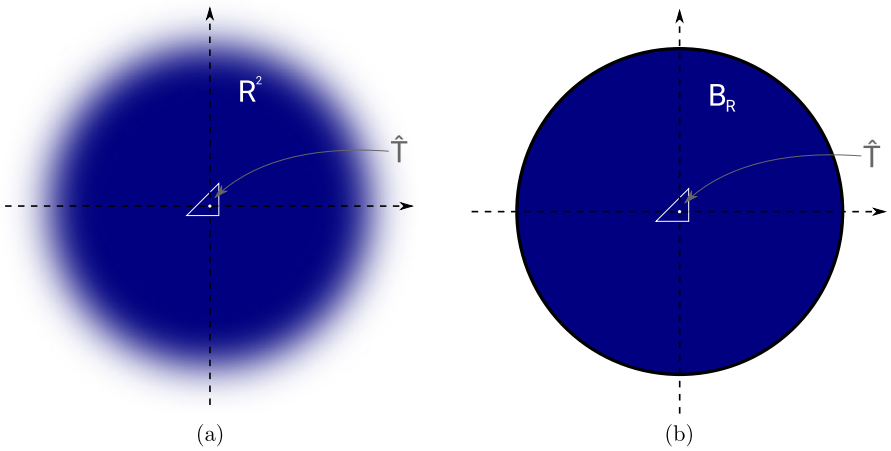


Fig. 8 **a** Unbounded domain with reference triangle $\hat{T}$. **b** Truncated domain $B_R(0)$ with reference triangle $\hat{T}$

and remark that, in the same way as pointed out in Remark 10, we can also precompute the matrices $\mathbf{\Gamma}_{\hat{T},\ell}$ for a four-dimensional array of values in an offline stage and interpolate them efficiently in the online stage. This way, we get the separable model

$$\hat{\mathcal{J}}_{\mathrm{SMWapprox}}(\boldsymbol{\eta}) := \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top \left(\boldsymbol{I}_2 - (\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)\mathbf{\Gamma}_{\hat{T},\ell}\right)^{-1} \nabla u_h|_{T_\ell} \tag{60}$$

as an approximation to the separable exact model (32).

6 Relationships

We investigate the relationship between the model (52) of Sect. 4 that is motivated by the continuous concept of topological derivatives and the model (60) of Sect. 5 which is meant to approximate the Sherman–Morrison–Woodbury matrix identity model (32). We restrict our presentation to the case of homogeneous material distribution around the fixed element $T_\ell \in \mathcal{T}$.

We start by recalling the definitions of the discretized weak polarization matrix $\mathcal{P}_{\hat{T},h}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}]$ (49) and the discretization of the matrix $\mathbf{\Gamma}_{\hat{T}}$ (59) in the homogeneous setting,

$$\mathcal{P}_{\hat{T},h}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}] = \left[\frac{1}{|\hat{T}|} \int_{\hat{T}} \nabla K_{\hat{T},h}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}; \boldsymbol{e}^{(1)}] \mathrm{d}x \quad \frac{1}{|\hat{T}|} \int_{\hat{T}} \nabla K_{\hat{T},h}[\lambda^{\mathrm{out}}, \lambda^{\mathrm{in}}; \boldsymbol{e}^{(2)}] \mathrm{d}x \right] \tag{61}$$

$$\mathbf{\Gamma}_{\hat{T},h}[\lambda^{\mathrm{out}}] = \left[\frac{1}{\hat{T}} \int_{\hat{T}} \nabla W_{\hat{T},h}[\lambda^{\mathrm{out}}; \boldsymbol{e}^{(1)}] \mathrm{d}x \quad \frac{1}{\hat{T}} \int_{\hat{T}} \nabla W_{\hat{T},h}[\lambda^{\mathrm{out}}; \boldsymbol{e}^{(2)}] \mathrm{d}x \right] \tag{62}$$

where $K_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}; \boldsymbol{e}^{(k)}] \in V_h^R$ is the solution to (48) and $W_{\hat{T},h}[\lambda^{\text{out}}; \boldsymbol{e}^{(k)}] \in V_h^R$ is the finite element approximation to (57), i.e.,

$$\int_{B_R(0)} \lambda_{\hat{T}} \nabla K_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}; \boldsymbol{e}^{(k)}] \cdot \nabla \psi_h \, \mathrm{d}x = -(\lambda^{\text{in}} - \lambda^{\text{out}}) \int_{\hat{T}} \boldsymbol{e}^{(k)} \cdot \nabla \psi_h \, \mathrm{d}x \tag{63}$$

$$\int_{B_R(0)} \lambda^{\text{out}} \nabla W_{\hat{T},h}[\lambda^{\text{out}}; \boldsymbol{e}^{(k)}] \cdot \nabla \psi_h \, \mathrm{d}x = -\int_{\hat{T}} \boldsymbol{e}^{(k)} \cdot \nabla \psi_h \, \mathrm{d}x \tag{64}$$

for all $\psi_h \in V_h^R$. Here, recall that $\lambda_{\hat{T}}(x) = \chi_{\hat{T}}(x)\lambda^{\text{in}} + \chi_{B_R(0)\setminus\hat{T}}(x)\lambda^{\text{out}}$.

We show the following relation between $\mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}]$ and $\boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}]$:

Theorem 9 *It holds that*

$$\mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}] = (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}] \left(\boldsymbol{I}_2 - (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}] \right)^{-1} \tag{65}$$

and further

$$\boldsymbol{I}_2 + \mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}] = (\boldsymbol{I}_2 - (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}])^{-1}. \tag{66}$$

Proof Recall that we use piecewise linear and globally continuous finite elements on a triangular mesh of M elements of $B_R(0)$ where, according to Remark 9, the subdomain $\hat{T}$ is resolved by exactly one triangle of the mesh. Let now the finite element stiffness matrix of (63) be denoted by $\tilde{\boldsymbol{A}}$ and the one of (64) by $\boldsymbol{A}$ where we use the same mesh and finite element space for both equations. Let $\hat{\ell} \in \{1, \ldots, M\}$ be the element index corresponding to the triangle $\hat{T}$. Note that the material distribution in (63) differs from that in (64) only in element $\hat{\ell}$ and we have $\boldsymbol{A} = \sum_{k=1}^{M} \lambda^{\text{out}} B_k B_k^\top$ and

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} + (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{B}_{\hat{\ell}} \boldsymbol{B}_{\hat{\ell}}{}^\top,$$

thus, an application of the Sherman–Morrison–Woodbury formula of Lemma 4 yields

$$\tilde{\boldsymbol{A}}^{-1} = \boldsymbol{A}^{-1} - (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{A}^{-1} \boldsymbol{B}_{\hat{\ell}} \left(\boldsymbol{I}_2 + (\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{B}_{\hat{\ell}}{}^\top \boldsymbol{A}^{-1} \boldsymbol{B}_{\hat{\ell}} \right)^{-1} \boldsymbol{B}_{\hat{\ell}}{}^\top \boldsymbol{A}^{-1}. \tag{67}$$

On the other hand, we know from Lemma 8 that for the chosen piecewise linear finite elements where $\hat{T}$ is resolved by only one triangle (i.e. $\nabla W_{\hat{T},h}[\lambda^{\text{out}}; \boldsymbol{e}^{(k)}](x)$, $\nabla K_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}; \boldsymbol{e}^{(k)}](x)$ are constant on $\hat{T}$), we have

$$(\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}] = -(\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{B}_{\hat{\ell}}{}^\top \boldsymbol{A}^{-1} \boldsymbol{B}_{\hat{\ell}}$$

$$\text{and} \quad \mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}] = -(\lambda^{\text{in}} - \lambda^{\text{out}}) \boldsymbol{B}_{\hat{\ell}}{}^\top \tilde{\boldsymbol{A}}^{-1} \boldsymbol{B}_{\hat{\ell}}.$$

Thus, denoting $\boldsymbol{\Gamma}^{\lambda}_{\hat{T},h} := (\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}]$ and plugging in (67) yields

$$\begin{aligned}\mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}] =& \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h} + \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h}(\boldsymbol{I}_2 - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1}\boldsymbol{\Gamma}^{\lambda}_{\hat{T},h} \\ =& \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h}(\boldsymbol{I}_2 + (\boldsymbol{I}_2 - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1}\boldsymbol{\Gamma}^{\lambda}_{\hat{T},h}) \\ =& \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h}(\boldsymbol{I}_2 - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1}\end{aligned}$$

where we used the identity $(\boldsymbol{I} - \boldsymbol{B})^{-1} = \boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{B})^{-1}\boldsymbol{B}$ for any matrix $\boldsymbol{B}$ such that $\boldsymbol{I} - \boldsymbol{B}$ is invertible in the last step. This proves (65). In order to see (66) note that, by (65), it holds $\boldsymbol{I}_2 + \mathcal{P}_{\hat{T},h}[\lambda^{\text{out}}, \lambda^{\text{in}}] = \boldsymbol{I}_2 + \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h}(\boldsymbol{I}_2 - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1} = (\boldsymbol{I}_2 - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h} + \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})(\boldsymbol{I} - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1} = (\boldsymbol{I} - \boldsymbol{\Gamma}^{\lambda}_{\hat{T},h})^{-1}$. □

Corollary 10 *From Theorem 9 it follows immediately that the two models $\hat{\mathcal{J}}_{TDnum}$ defined in (52) and $\hat{\mathcal{J}}_{SMWapprox}$ defined in (60) coincide.*

Remark 11 We remark that the same proof can be conducted in the case of inhomogeneous material around the element of interest and the statements of Theorem 9 and Corollary 10 remain valid also in that case.

At the first glance, Theorem 9 and Corollary 10 seem very surprising since they state that the model that is based on the continuous concept of topological derivatives brought to a discrete setting coincides with a model that is based on a certain approximation of a term generated by the purely algebraic Sherman–Morrison–Woodbury matrix identity. This resemblance, however, has been identified in [32] on the purely continuous setting for the case of elliptic inclusions.

Lemma 11 ([32]) *Assume that ω is an ellipse. Then*

$$\mathcal{P}_{\omega}[\lambda^{\text{out}}, \lambda^{\text{in}}] = -(\lambda^{\text{in}} - \lambda^{\text{out}})\left(\boldsymbol{I}_2 + (\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Psi}[\lambda^{\text{out}}]\right)^{-1}\boldsymbol{\Psi}[\lambda^{\text{out}}] \tag{68}$$

where $\boldsymbol{\Psi}[\lambda^{\text{out}}] \in \mathbb{R}^{2\times 2}$ is given by

$$\boldsymbol{\Psi}[\lambda^{\text{out}}]_{i,j} = -\left(\int_{\partial\omega} n(x)\nabla_x \Phi[\lambda^{\text{out}}](x)^{\top}\, ds_x\right)_{i,j} = -\int_{\partial\omega} n_i \partial_{x_j}\Phi[\lambda^{\text{out}}](x)\, ds_x \tag{69}$$

with the fundamental solution $\Phi[\lambda^{\text{out}}]$ of the operator $u \mapsto -div(\lambda^{\text{out}}\nabla u)$, i.e., $\Phi[\lambda^{\text{out}}](x) = -1/(2\pi\lambda^{\text{out}})ln(|x|)$.

Proof This follows straightforwardly from [32, Sec. 8] by restricting the (vector-valued) elasticity problem treated there to the scalar Laplace-type problem considered here. □

From Lemma 11, it follows in the same way as in the proof of Theorem 9 that

$$\boldsymbol{I}_2 + \mathcal{P}_{\omega}[\lambda^{\text{out}}, \lambda^{\text{in}}] = (\boldsymbol{I}_2 + (\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Psi}[\lambda^{\text{out}}])^{-1}, \tag{70}$$

and thus, from (21), we get the alternative representation of the topological derivative for elliptic inclusion shapes ω

$$d\mathcal{J}[\lambda^{\text{out}}, \lambda^{\text{in}}](\Omega)(z, \omega) = (\lambda^{\text{in}} - \lambda^{\text{out}})\nabla u(z)^{\top}(\boldsymbol{I}_2 + (\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Psi}[\lambda^{\text{out}}])^{-1}\nabla p(z). \tag{71}$$

Thus, the matrix $\boldsymbol{\Gamma}_{\hat{T},h}[\lambda^{\text{out}}]$ used in $\hat{\mathcal{J}}_{\text{SMWapprox}}$ can also be seen as an approximation to the negative fundamental matrix $-\boldsymbol{\Psi}[\lambda^{\text{out}}]$. Finally, note that the proof of Lemma 11 in [32] is not valid for triangular inclusion shapes, however, at the discrete level, relation (65) still holds. Furthermore, note that the matrix $\boldsymbol{\Psi}[\lambda^{\text{out}}]$ is symmetric such that the right hand side of (68) can also be written as $-(\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Psi}[\lambda^{\text{out}}]\left(\boldsymbol{I}_2 + (\lambda^{\text{in}} - \lambda^{\text{out}})\boldsymbol{\Psi}[\lambda^{\text{out}}]\right)^{-1}$ which coincides with the structure of (65).

7 Numerical Experiments

In this section, we examine the models introduced in Sects. 4 and 5 and compare them to the exact solution as well as the diagonal approximation model introduced in Sect. 3. Since we noted in Sect. 6 that model (52) of Sect. 4 and model (60) of Sect. 5 coincide, we will here only consider the latter model. We remark that coincidence of the two models was observed also in all numerical examples.

All numerical results are illustrated for the model problem introduced in Sect. 2.1 with the two-dimensional computational domain $\mathsf{D} = (0, 1)^2$ with Dirichlet and Neumann boundaries $\Gamma_D = \{(0, y), y \in (0, 1)\} \cup \{(x, 0), x \in (0, 1)\}$, $\Gamma_N = \partial\mathsf{D}\backslash\Gamma_N$ with corresponding data $g_D = 0$, $g_N(x_1, x_2) = x_1 x_2$ and the constant source term $f(x_1, x_2) = 1$. The material coefficient $\lambda(x)$ will vary between the values $\underline{\lambda} = 1$ and $\overline{\lambda} = 1000$.

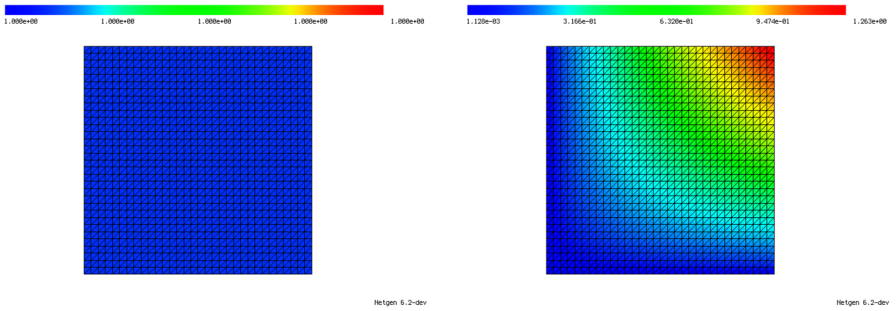
We begin by considering a homogeneous setting.

7.1 Homogeneous Material Distribution

Here, we consider a constant material distribution, i.e., $\lambda_\Omega(x) = \lambda^{\text{out}}$ for all $x \in \mathsf{D}$, which corresponds to setting $\Omega = \emptyset$ for some value λ^{out} in the setting of Sect. 2.1.1. See also Fig. 9 for plots of the material distribution and the finite element solution on a mesh with $n = 1089$ nodes and $m = 2048$ elements.

7.1.1 Numerical Comparison of Different Models for Fixed Element

We fix the interior element $T_\ell \in \mathcal{T}$ as that element of type 1 (cf. Fig. 1) that has the point $(0.5, 0.25)$ as its bottom right vertex and compare the different models we introduced in the previous sections for the case where the homogeneous material distribution is perturbed only in that one element T_ℓ, i.e., $\boldsymbol{\eta} = \boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}$. Figure 10 shows the different models as functions of the perturbed value $\eta \in [\underline{\lambda}, \overline{\lambda}] = [1, 1000]$ for three different background material values $\lambda^{\text{out}} = 1$, $\lambda^{\text{out}} \approx 145.834$, $\lambda^{\text{out}} = 1000$



(a) (b)

Fig. 9 **a** Material coefficient $\lambda(x)$ for homogeneous setting. **b** Finite element solution u_h of problem (5b) with data specified in Sect. 7 for homogeneous material distribution

(cf. Remark 12). In Fig. 10 we can see the exact solution $\mathcal{J}(\boldsymbol{\eta})$ (where the perturbed stiffness matrix is inverted) for certain values of η along with the exact Sherman–Morrison–Woodbury model $\hat{\mathcal{J}}_{\text{SMW}}$ (32) which shows, as expected, perfect coincidence. Moreover, we can see the diagonal approximation of the Sherman–Morrison–Woodbury model $\hat{\mathcal{J}}_{\text{SMWdiag}}$ (35) which shows a certain error, but can be evaluated more efficiently. The models $\hat{\mathcal{J}}_{\text{TDnum}}$ (52) and $\hat{\mathcal{J}}_{\text{SMWapprox}}$ (60) can be seen to match exactly, as was predicted by Theorem 9 and Corollary 10. Moreover, it can be seen from Fig. 10 that these two models approximate the exact solution extraordinarily well while being cheap to evaluate during the online stage (after precomputations have been done in an offline stage, cf. Remark 10). For comparison, we also included the linearization $\mathcal{J}(\boldsymbol{\eta}) \approx \mathcal{J}(\boldsymbol{\lambda}) - |T_\ell|(\eta - \boldsymbol{\lambda}_\ell)|\nabla u_h|_{T_\ell}|^2$ and the topological derivative model when the analytical formula for the topological derivative of circular inclusions (22) is used. It can be seen that the linearization model is far away from the true solution. But also the latter model shows a significant error which confirms the necessity to account for the triangular inclusion shape as it was done in Sect. 4.

In order to quantify these errors, let us define the relative error measure of a model $\hat{\mathcal{J}}$ in an element T_ℓ for a given material distribution $\boldsymbol{\lambda}$ by

$$\delta\hat{\mathcal{J}}[T_\ell] := \frac{1}{\Delta\mathcal{J}[T_\ell]} \max_{\eta\in[\underline{\lambda},\overline{\lambda}]} \left|\hat{\mathcal{J}}\left(\boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}\right) - \mathcal{J}\left(\boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}\right)\right| \quad (72)$$

where $\Delta\mathcal{J}[T_\ell] = \max_{t\in(\underline{\lambda},\overline{\lambda})}\mathcal{J}(\boldsymbol{\lambda} + (t - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)}) - \min_{t\in(\underline{\lambda},\overline{\lambda})}\mathcal{J}(\boldsymbol{\lambda} + (t - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)})$ is the difference of maximal and minimal values of the exact model in T_ℓ. Thus, $\delta\hat{\mathcal{J}}[T_\ell]$ measures the maximum relative error of a model $\hat{\mathcal{J}}$ in element T_ℓ relative to the variation of the exact cost function $\mathcal{J}$. The relative errors according to (72) for the three expansion points $\lambda(x) = \lambda^{\text{out}}$ investigated in Fig. 10 are as follows: For the linearization model the errors are as high as (20 954%, 85.61%, 25.21%), for the model $\hat{\mathcal{J}}_{\text{SMWdiag}}$ they are (16.65%, 3.41%, 0.99%), for $\hat{\mathcal{J}}_{\text{TDcirc}}$ we have (57.93%, 27.23%, 49.43%) and for the coinciding models $\hat{\mathcal{J}}_{\text{TDnum}}$ and $\hat{\mathcal{J}}_{\text{SMWapprox}}$ we have the values (1.18%, 0.74%, 1.46%).

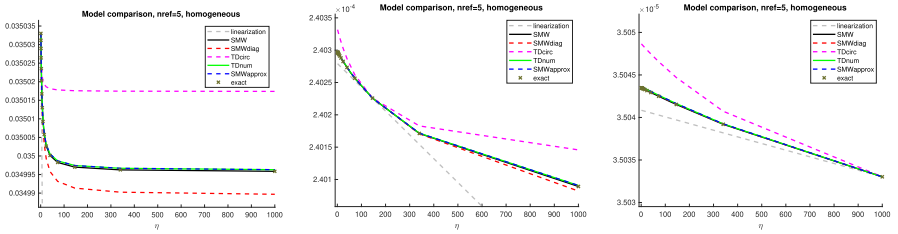


Fig. 10 Comparison of different models for homogeneous setting with background material **a** $\lambda^{\text{out}} = 1$, **b** $\lambda^{\text{out}} \approx 145.834$, **c** $\lambda^{\text{out}} = 1000$ as functions of the perturbed material coefficient in a fixed element T_ℓ

Table 1 Values $\eta^{(k)}$, $k = 1, \ldots, 16$, used for visualization and precomputation

1	1.252	1.590	2.050	2.688	3.596	4.921	6.917
10.035	15.127	23.901	40.072	72.563	145.834	340.187	1000

Remark 12 In Fig. 10, the models as well as the exact solution were evaluated at certain perturbation values $\eta^{(k)} \in [1, 1000]$, $k = 1, \ldots, 16$. These values have been chosen in the following way: It was observed numerically that the exact solution in Fig. 10 behaves similarly to $a + b\eta^{-0.5}$ for some constants a, b. Based on this observation, the points $\eta^{(k)}$ were chosen in such a way that, using these points as interpolation nodes, a piecewise linear interpolation of $a + b\eta^{-0.5}$ yields an equilibrated error, see Fig. 11. This was achieved by solving a system of nonlinear equations ensuring that the maximum interpolation error between any two neighboring nodes is equal. The values for $N = 16$ points are given in Table 1. While these points here are used solely for visualization purposes, their role will become more important in Sect. 7.2 to decide for which values of the material coefficient the (computationally expensive) precomputation should be carried out.

Remark 13 Also the choice of the Hölder parameter $\alpha = -0.5$ in (50) was motivated by the observation that the exact solution in Fig. 10a behaves roughly like $a+b\eta^{-0.5} =: f(\eta)$. Thus, an effective value for λ can be obtained by the relation $f(\lambda) = \sum w_i f(\lambda_i)$ for given values λ_i with corresponding weights w_i such that $\sum w_i = 1$, which, by inverting f, results in the chosen average value (50).

7.1.2 Numerical Comparison of Different Models on Whole Domain

In Fig. 10, we compared several different models for a fixed triangle in the mesh when the material perturbation parameter is varied between $\underline{\lambda} = 1$ and $\overline{\lambda} = 1000$. Next, we investigate the maximum relative error $\delta\hat{\mathcal{J}}$ as defined in (72) of a model $\hat{\mathcal{J}}$ as a function of the position in space.

Given the findings of Fig. 10, we will focus on the two approximate Sherman–Morrison–Woodbury models $\hat{\mathcal{J}}_{\text{SMWdiag}}$ and $\hat{\mathcal{J}}_{\text{SMWapprox}}$ (recall that $\hat{\mathcal{J}}_{\text{TDnum}}$ coincides with $\hat{\mathcal{J}}_{\text{SMWapprox}}$). In Fig. 12, we plot the maximum relative errors (72) for these two

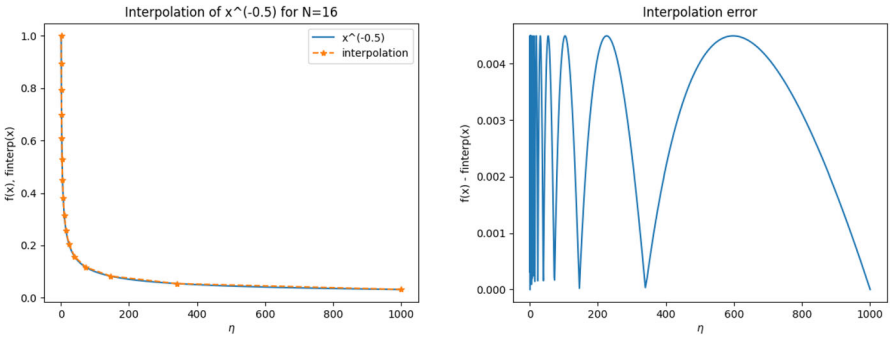


Fig. 11 The interpolation nodes $\eta^{(k)}$, $k = 1, \ldots, 16$ are chosen in such a way that the maximum interpolation error between the function $x^{-0.5}$ and its piecewise linear interpolant attains the same value in each interval $(\eta^{(k)}, \eta^{(k+1)})$

models for all interior elements of the computational domain for the homogeneous material distribution $\lambda(x) = \lambda^{\text{out}} = 1$, i.e., we are in the setting of Fig. 10a. Elements touching the boundary are discussed separately in Sect. 7.3.1. Here, we can see that the maximum relative error of the model $\hat{\mathcal{J}}_{\text{SMWdiag}}$ is around 47% whereas it is around 17% for $\hat{\mathcal{J}}_{\text{SMWapprox}}$. The four different rows of Fig. 12 show different threshold values for the relative errors in the color bars. Moreover, it can be seen from the right column in Fig. 12 that the model $\hat{\mathcal{J}}_{\text{SMWapprox}}$ behaves particularly well in the center of the homogeneous domain and that the error increases slightly the closer one gets to a boundary. This was to be expected since the idea of the model $\hat{\mathcal{J}}_{\text{SMWapprox}}$ in Sect. 5 (and equivalently model $\hat{\mathcal{J}}_{\text{TDnum}}$ of Sect. 2.2) was to zoom in locally around the fixed element T_ℓ and assume that all boundaries are sufficiently far away, cf. Figs. 7 and 8. Nevertheless, the maximum error attained by $\hat{\mathcal{J}}_{\text{SMWapprox}}$ is still significantly smaller than that of the model $\hat{\mathcal{J}}_{\text{SMWdiag}}$.

7.1.3 Towards Topology Optimization Using Approximate Models

We want to illustrate the potential of the introduced models $\hat{\mathcal{J}}_{\text{SMWdiag}}$ and $\hat{\mathcal{J}}_{\text{SMWapprox}}$ (which coincides with $\hat{\mathcal{J}}_{\text{TDnum}}$) in the course of a binary topology optimization algorithm. Here, we simply decide for each element T_ℓ if it should be occupied by $\lambda^{\text{out}} = 1$ or $\lambda^{\text{in}} = 1000$ based on the values of a model $\hat{\mathcal{J}}(\boldsymbol{\lambda} + (\eta - \boldsymbol{\lambda}_\ell)\boldsymbol{e}^{(\ell)})$ at $\eta = \lambda^{\text{out}}$ and $\eta = \lambda^{\text{in}}$, i.e., we do not allow for intermediate material values. Here, we again start out from the homogeneous design where $\boldsymbol{\lambda} \in \mathbb{R}^m$ is the constant one vector. As a reference, we consider the separable exact model $\hat{\mathcal{J}}_{\text{SMW}}$ (32). Due to this model's properties, the material distribution obtained by the mentioned procedure is a local minimum, which cannot be improved by switching the state of only one element. Note that this is a stronger notion of optimality than a design being solely a stationary point of the relaxed optimization problem. From a theoretical point of view it is not entirely clear that using approximations of exact separable models such as $\hat{\mathcal{J}}_{\text{SMWdiag}}$ the same effect can be achieved. However, already in [19] it was reported that the SGP concept combined with an approximation of the $\hat{\mathcal{J}}_{\text{SMWdiag}}$ type lead to a much better

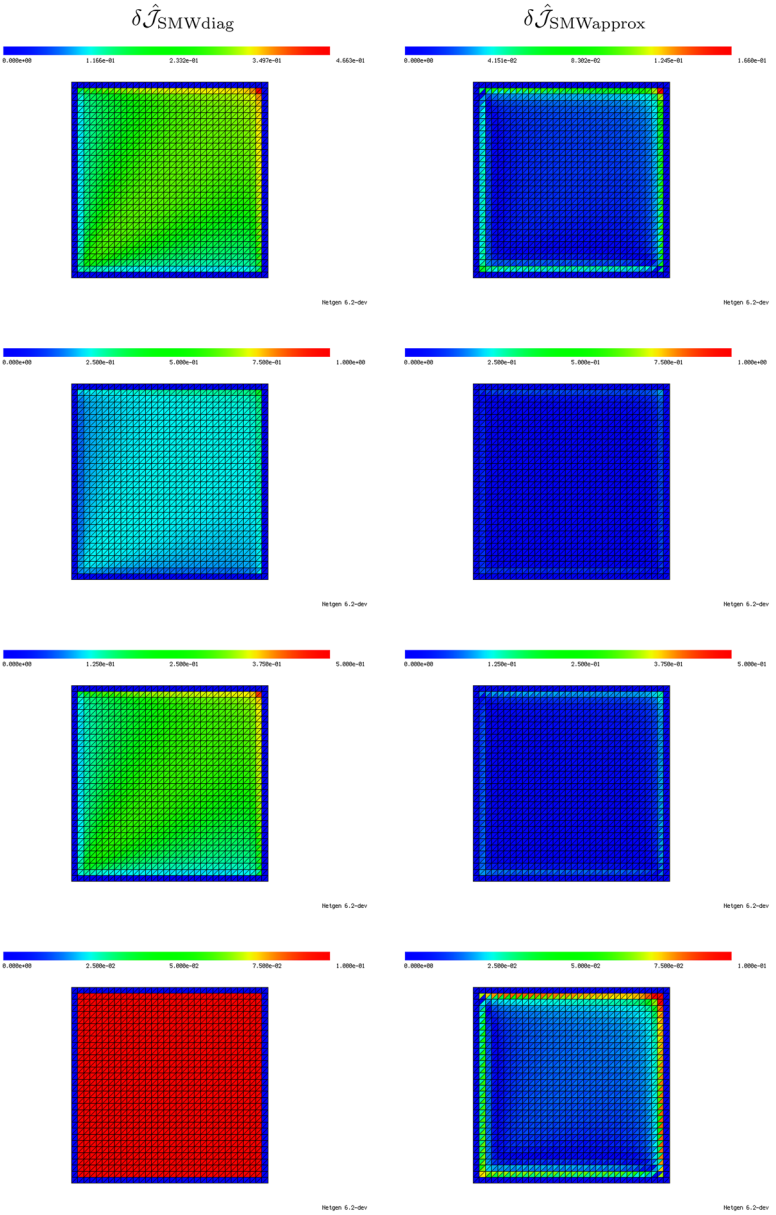


Fig. 12 Comparison of relative errors $\delta\hat{\mathcal{J}}[T_\ell]$ according to (72) for models $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ (35) in left column and $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ (60) (= $\hat{\mathcal{J}}_{\mathrm{TDnum}}$ (52)) in right column for all interior elements T_ℓ in homogeneous setting. First line shows color plot according to their respective maximum errors. Second to fourth line show threshold for maximum relative error at 100%, 50% and 10%, respectively. Errors in elements touching the boundary are not computed

local minimizer for a binary topology optimization problem than the MMA method utilizing convex separable approximations. Here, we investigate this effect in more detail using a selection of the previously suggested models. In addition, we also make a comparison with an MMA model

$$\begin{aligned}\hat{\mathcal{J}}_{\mathrm{MMA}}(\boldsymbol{\eta}) &= \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\nabla u_h|_{T_\ell})^\top \left(\boldsymbol{I}_2 - \frac{\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell}{L - \boldsymbol{\lambda}_\ell}\boldsymbol{I}_2\right)^{-1} \nabla u_h|_{T_\ell} \\ &= \mathcal{J}(\boldsymbol{\lambda}) - \sum_{\ell=1}^{m} |T_\ell||\nabla u_h|_{T_\ell}|^2 \frac{(\boldsymbol{\eta}_\ell - \boldsymbol{\lambda}_\ell)(\boldsymbol{\lambda}_\ell - L)}{(\boldsymbol{\eta}_\ell - L)}. \end{aligned} \tag{73}$$

Here L plays the role of a vertical asymptote, which is chosen individually for each element by a heuristic update scheme in the original MMA method, see [17]. As we consider only a single update step here, the heuristic for the choice of L can not be applied. Instead we test three different constant choices of the asymptote, $L = 0$, $L = -5$, $L = -10$. Note that (73) can be obtained from (34) by replacing $\boldsymbol{\Gamma}^{(\ell)}$ by $\frac{1}{L-\boldsymbol{\lambda}_\ell}\boldsymbol{I}_2$.

Since, as it is well-known, the optimum material design for compliance minimization without limitation on the volume is the full design, we here include a simple volume penalization in the cost function and use the augmented cost function

$$\mathcal{L}(\boldsymbol{\lambda}) := \mathcal{J}(\boldsymbol{\lambda}) + \omega \mathrm{Vol}(\boldsymbol{\lambda}) \tag{74}$$

with a fixed weight $\omega = 7.5$ and the volume of the strong material $\mathrm{Vol}(\boldsymbol{\lambda}) = \sum_{\ell=1}^{m} |T_\ell|(\boldsymbol{\lambda}_\ell - \lambda^{\mathrm{out}})/(\lambda^{\mathrm{in}} - \lambda^{\mathrm{out}})$. Note that $\mathrm{Vol}(\boldsymbol{\lambda})$ itself is a separable function which can be dealt with without approximation error. Figure 13 shows the designs obtained after one step of the procedure mentioned above when using (a) the exact (but expensive) model $\hat{\mathcal{J}}_{\mathrm{SMW}}$, (b) the diagonal approximation $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$, (c) the proposed model $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ and (d)–(f) the MMA model with $L = 0$, $L = -5$ and $L = -10$. Comparing pictures (b) and (c) to (a), we see that the error in the design produced by the model in (c) is almost zero, whereas it is a bit larger for the diagonal approximation model in (b). The performance of the MMA model here depends heavily on the choice of the parameter L. For illustration of the method, we also plotted the curves corresponding to the exact and the two mentioned approximate models in three fixed elements. Figure 13g shows that, in the leftmost of the three highlighted elements in (a)–(f), the value of all six considered models at $\lambda^{\mathrm{in}} = 1000$ is higher than at $\lambda^{\mathrm{out}} = 1$, thus making a switching of the material from λ^{out} to λ^{in} unattractive. In the same way, in the rightmost of the three marked elements, all models except for MMA with $L = 0$ show smaller values at λ^{in} than at λ^{out}, thus suggesting switching the material to decrease the cost function, see Fig. 13i. In the central one out of these three elements, however, the diagonal approximation model $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ suggests to switch the material since its value is smaller at λ^{in} than at λ^{out}, whereas the exact model as well as the proposed approximation $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ suggest not to switch it, see Fig. 13h. The MMA model shows good behavior for the choice $L = -5$, but large errors for the choices $L = 0$ and $L = -10$. Finally, we remark that, when comparing

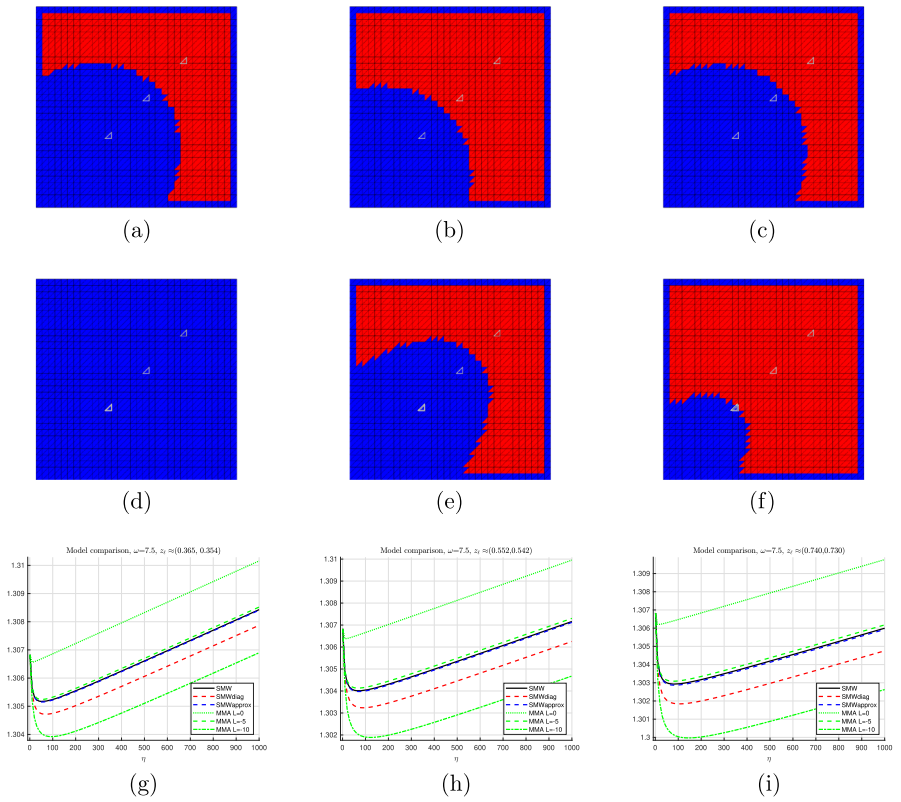
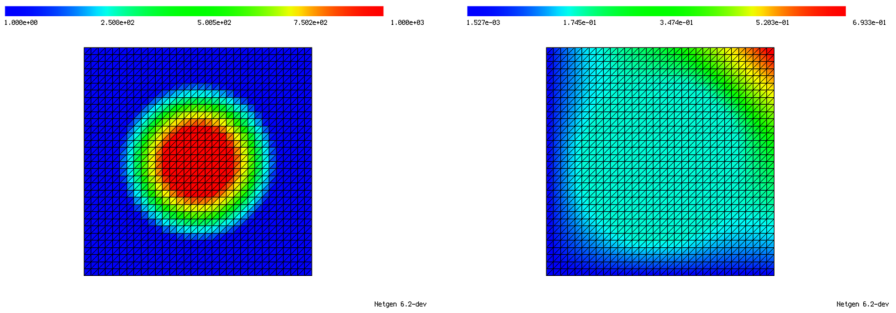


Fig. 13 Top and central row: Material distribution after one step of binary topology optimization for (74) with $\omega = 7.5$ starting out from homogeneous material $\lambda(x) = 1$ (see Fig. 9a) when using **a** exact compliance model $\hat{\mathcal{J}}_{\mathrm{SMW}}$, **b** diagonal approximation to Sherman–Morrison–Woodbury model $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$, **c** approximate compliance model $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$, **d** MMA model with $L = 0$, **e** MMA model with $L = -5$, **f** MMA model with $L = -10$. Bottom row: Illustration of local models in three triangles marked in top row from bottom left (**g**) to top right (**i**)

with the exact model $\hat{\mathcal{J}}_{\mathrm{SMW}}$, for $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ the wrong decision was taken for 280 out of 1800 interior elements whereas this was the case only for 19 elements in the case of $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$. For the MMA model with $L = 0$, $L = -5$, $L = -10$, the numbers of wrongly switched elements were 930, 102 and 586 elements, respectively.

7.2 Inhomogeneous Material Distribution

Next we consider a numerical example with an inhomogeneous material distribution as it may appear in the course of a density-based topology optimization algorithm, which is the motivation for this work. We consider a material coefficient $\lambda(x)$ that continuously varies between $\underline{\lambda} = 1$ and $\overline{\lambda} = 1000$ as



(a) (b)

Fig. 14 **a** Material coefficient λ for inhomogeneous setting. **b** Finite element solution u_h of problem (5b) with data specified in Sect. 7 for inhomogeneous material distribution

$$\lambda(x) = \begin{cases} \underline{\lambda}, & |x-m| \geq r_2, \\ \underline{\lambda} + \frac{|x-m|-r_1}{r_2-r_1}(\overline{\lambda} - \underline{\lambda}), & |x-m| \in (r_1, r_2), \\ \overline{\lambda}, & |x-m| \leq r_1, \end{cases}$$

with $r_1 = 0.15$, $r_2 = 0.35$ and $m = (0.5, 0.5)^\top$. The material distribution and the corresponding finite element solution of (5b) with the data defined in the beginning of Sect. 7 are depicted in Fig. 14.

7.2.1 Numerical Evaluation of $\hat{\mathcal{J}}_{\text{SMWapprox}}$ in Inhomogeneous Setting

In order to evaluate $\hat{\mathcal{J}}_{\text{SMWapprox}}$ in inhomogeneous regions of the computational domain, recall that model (60) involves the matrix $\mathbf{\Gamma}_{\hat{T},\ell} = \mathbf{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})]$ defined in (59) where $\lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}$ are averaged material coefficients according to (50), see also Fig. 6. The matrix $\mathbf{\Gamma}_{\hat{T},\ell}$, in turn, is based on the solution to the truncated transmission problem (58) for the given averaged material values. Thus, in order to evaluate $\hat{\mathcal{J}}_{\text{SMWapprox}}$, that exterior problem would have to be solved for the averaged material values of every single element, which would make the model computationally intractable.

For that reason, we introduce another approximation step: We precompute the matrix $\mathbf{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})]$ for a large discrete set of combinations of material coefficients in an offline stage. More precisely, we precompute the matrix $\mathbf{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})]$ for all combinations of tuples $(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}) \in \{\eta^{(1)}, \ldots, \eta^{(N)}\}^4$ where $\underline{\lambda} = \eta^{(1)} < \cdots < \eta^{(N)} = \overline{\lambda}$ and for the two types of elements $\hat{T} = \hat{T}^{(1)}$ and $\hat{T} = \hat{T}^{(2)}$, cf. Fig. 1. Since (58) has to be solved for $k = 1, 2$, this yields a total of $4N^4$ finite element solutions of truncated transmission problems. In the online stage, given averaged values $(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3}) \notin \{\eta^{(1)}, \ldots, \eta^{(N)}\}^4$, the matrix $\mathbf{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})]$ is approximated by piecewise linear interpolation of the precomputed data.

In our experiments, in order to numerically approximate (58), we used the moderately large value $R = 30$ for the radius of the computational domain and discretized it by a mesh consisting of about 4400 triangular elements and about 2300 vertices. We chose $N = 16$ material points between $\underline{\lambda} = 1$ and $\overline{\lambda} = 1000$ which were chosen as stated in Table 1. Since increasing the number of material points N will drastically increase the precomputation time, the concrete choice of these points is of big importance. Thus, the points were chosen such that the interpolation error that is made in the online stage is as small as possible, see Remark 12. The total precomputation time for this setting was about two hours on a single core. Note that, for given PDE constraint, discretization method and material catalogue $\{\eta^{(1)}, \dots, \eta^{(N)}\}$, this precomputation step has to be performed only once and can henceforth be used in all optimization runs.

7.2.2 Numerical Comparison of Models on Computational Domain

We make the same comparison of the two most promising models $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ (35) and $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ (60) as it was done for the homogeneous setting in Sect. 7.1.2. Again, recall that $\hat{\mathcal{J}}_{\mathrm{TDnum}}$ coincides with $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ and is thus not examined separately.

Figure 15 again shows the maximum relative error $\delta\hat{\mathcal{J}}$ of these two models over the computational domain. Again, different thresholds of the color bar are shown. As it was already observed in Sect. 7.1.2, the model $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ behaves particularly well in regions of homogeneous material. For both models, the largest errors occur at the transition from homogeneous material $\lambda(x) = \underline{\lambda}$ to inhomogeneous material.

From Fig. 15 it can also be seen that the largest error of $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ is around 315% compared to only about 100% for $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$. However, we mention that this effect disappears when a finer mesh is chosen as it is illustrated in Fig. 16. There, it can be seen that the maximum error of $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ in the refined mesh is only around 86% which is in the same range as for $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$. The reason for this improvement of $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ is that, for a given function $\lambda(x)$, as the mesh size decreases, the range of values to be averaged in the direct neighborhood of an element becomes smaller which results in a smaller error when computing the average values $\lambda_{T_\ell}^{S_k}$ (50), $k = 1, 2, 3$. In general, the model $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ behaves well if material variations around a fixed element are small and makes larger approximation errors when large ranges of material values have to be averaged.

7.3 Further Improvements

Finally, we point out several directions in which this research could be extended to further improve the models $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$, $\hat{\mathcal{J}}_{\mathrm{TDnum}}$.

7.3.1 Boundary Regions

So far, we restricted our numerical results to regions away from the boundary and did not treat elements that touch the boundary. The reason for this is that, in the

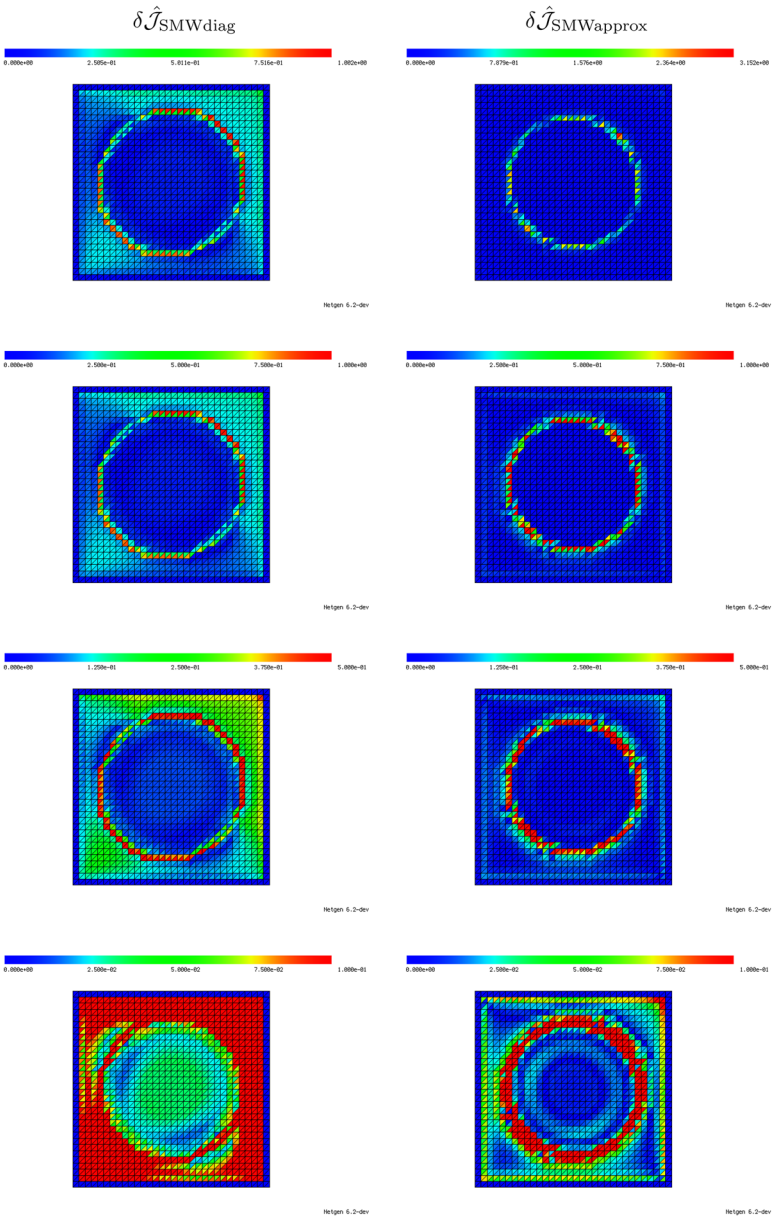


Fig. 15 Comparison of relative errors $\delta\hat{\mathcal{J}}[T_\ell]$ according to (72) for models $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ (35) in left column and $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$ (60) (= $\hat{\mathcal{J}}_{\mathrm{TDnum}}$ (52)) in right column for all interior elements T_ℓ in inhomogeneous setting. First line shows color plot according to their respective maximum errors. Second to fourth line show threshold for maximum relative error at 100%, 50% and 10%, respectively. Errors in elements touching the boundary are not computed

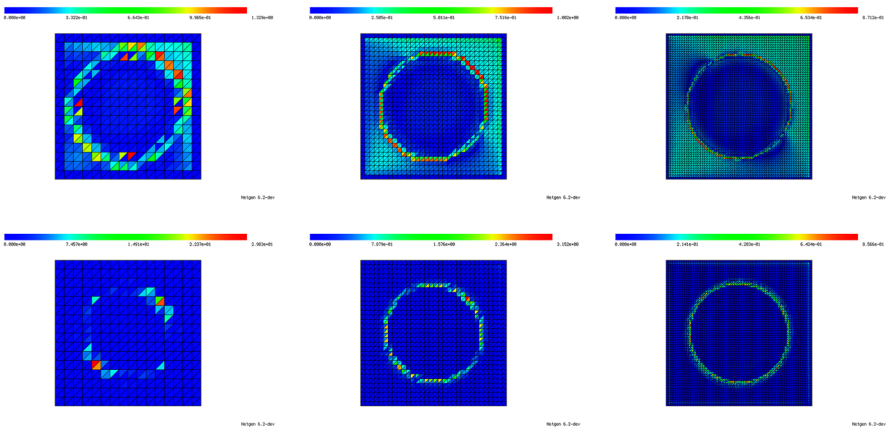


Fig. 16 Comparison of relative errors in inhomogeneous setting for different mesh sizes for models $\hat{\mathcal{J}}_{\mathrm{SMWdiag}}$ (top row) and $\hat{\mathcal{J}}_{\mathrm{SMWapprox}} = \hat{\mathcal{J}}_{\mathrm{TDnum}}$ (bottom row). Errors in elements touching the boundary are not computed

derivation of the models (52) and (60), the truncated exterior problems (48) and (58) are obtained by zooming in around the fixed element T_ℓ and rescaling. Thus, in the case where T_ℓ touches the boundary, the truncated domains depicted in Figs. 4 and 8 do not mimic the neighborhood of T_ℓ. Instead, it would be more appropriate to perform precomputations on truncated half spaces as depicted in Fig. 17. The figure shows the setting of problem (57) in the case of a homogeneous material distribution corresponding to elements touching the top boundary. Also, here, inhomogeneous material can be treated by precomputing a range of combinations of material values in an offline stage and interpolating averaged sector values in the online stage. Here, the precomputation becomes a bit more involved since, in addition to accounting for different materials and different element types and $k = 1, 2$ in (57), one also has to distinguish between a left, bottom, right or top boundary as well as between Dirichlet or Neumann conditions imposed on that boundary. Thus, in order to also treat boundary regions of a rectangular domain D, an additional $32N^4$ truncated half space problems have to be solved in the offline stage. Here, N is the number of used material values, e.g., $N = 16$.

In Fig. 18, we illustrate the improvement when elements touching the top boundary are given a special treatment by precomputing the matrix $\mathbf{\Gamma}_{\hat{T}}[(\boldsymbol{\lambda}_\ell, \lambda_{T_\ell}^{S_1}, \lambda_{T_\ell}^{S_2}, \lambda_{T_\ell}^{S_3})]$ also for the reference elements $\hat{T}$ depicted in Fig. 17. In addition to the data presented in Fig. 15, we also computed the maximum relative errors in all elements touching the top boundary (of Neumann type). If the same data as in the interior is used, the maximum relative error is attained in the elements at the boundary and is as high as 77.89%. When the mentioned treatment of the boundary elements is used, the maximum error is still attained at an interior element and is only 16.6%.

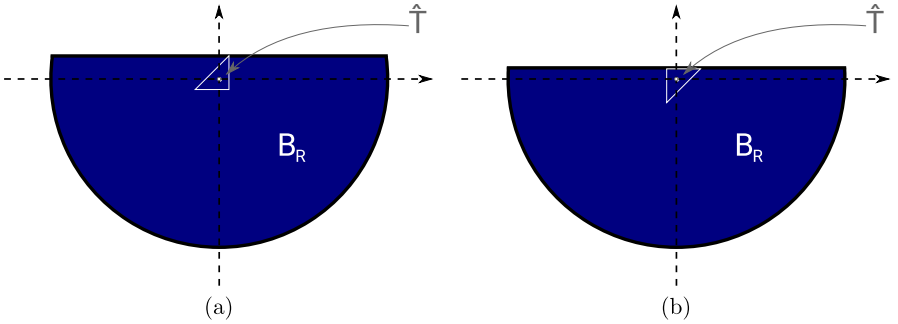


Fig. 17 Setting for exterior problem (48) corresponding to elements **a** of type 1 or **b** of type 2 that touch a top boundary. On the circular part of the boundary of B_R, homogeneous Dirichlet conditions are set. The boundary conditions at top can be either of Dirichlet or Neumann type, depending on the physical problem

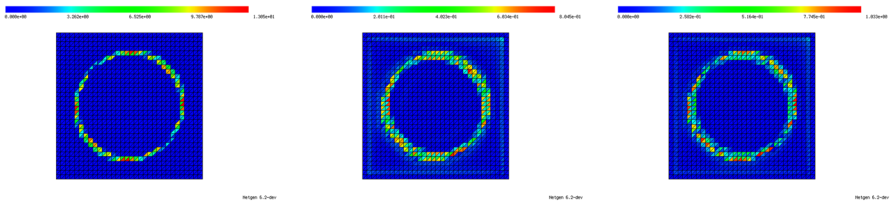
Fig. 18 Comparison of model $\hat{\mathcal{J}}_{\text{SMWapprox}}$ in Neumann boundary region **a** without and **b** with special precomputation using truncated half spaces as depicted in Fig. 17. For comparison, the same color scale that is cut off at 16.6% is used in (**a**) and (**b**). The maximum relative error in (**a**) is 77.89% whereas it is only 16.6% in (**b**)

7.3.2 Averaging of Inhomogeneous Material Distribution

We mention that we observed that the way material values are averaged over sectors has a strong impact on the obtained relative error. Figure 19 shows the same experiments as discussed in Sect. 7.2 for different values of α in the averaging process (50). For $\alpha = 1$, the weighted Hölder mean (50) reduces to the weighted arithmetic mean which yields a large maximum error of about 1300%, see Fig. 19(a). Recall that our choice $\alpha = -0.5$ yielded the result in the right column of Fig. 15 with maximal value of $\delta\hat{\mathcal{J}} \approx 315\%$. Further numerical studies for $\alpha = -0.2$ and $\alpha = -0.1$ are depicted in Fig. 19b,c showing that for $\alpha = -0.2$ the maximal error in the mesh is actually smaller than for the model $\hat{\mathcal{J}}_{\text{SMWdiag}}$.

Moreover, one might want to think of decomposing the domain $B_R(0)$ into more than three sectors in order to reduce the error made by the averaging of inhomogeneous material parameters. However, here one should keep in mind that the number of truncated exterior problems to be solved in the precomputation stage with n_{sec} sectors and N material points is of the order $N^{n_{\text{sec}}+1}$ and thus grows very fast with n_{sec}.

Finally, taking the average of material values on more than one layer of elements (cf. Fig. 6) together with suitable distance-dependent weights could lead to a better representation of the local material configuration and thus to potentially higher accuracy of the models $\hat{\mathcal{J}}_{\text{SMWdiag}}$ and $\hat{\mathcal{J}}_{\text{TDnum}}$ in an inhomogeneous setting.



(a) $\alpha = 1$ (b) $\alpha = -0.2$ (c) $\alpha = -0.1$

Fig. 19 Maximum relative error $\delta\hat{\mathcal{J}}_{\text{SMWapprox}}$ over computational domain for different averaging parameters α in (50)

Conclusion and Outlook

In this paper, we introduced and examined different separable approximations to a discretized topology/material optimization problem. The Sherman–Morrison–Woodbury formula applied to the perturbed finite element stiffness matrix yielded a first separable exact model which, however, is prohibitively expensive to evaluate. A diagonal approximation of the stiffness matrix yielded a first tractable model. We introduced a model that is motivated by the continuous concept of topological derivatives for triangular inclusion shapes. Moreover, we also introduced a model that approximates the Sherman–Morrison–Woodbury model with high accuracy by performing similar rescaling steps as in the topological derivative model. Subsequently, we showed the somewhat surprising result that these latter two models coincide. Finally, we compared the performances of all models numerically. While the diagonal approximation of the Sherman–Morrison–Woodbury model can be evaluated very efficiently without any problems, the new models need data to be precomputed in an offline stage. In our model problem, however, we saw that the newly introduced models show significantly higher accuracy in most regions of the domain.

This work presented here can be extended and continued in several directions.

- We illustrated our methods for the case of the compliance functional in a stationary heat equation. We emphasize that this model was chosen for compactness of presentation and that extensions to other cost functions and other linear PDE constraints (e.g., linear elasticity) can be obtained in a rather straight-forward way (possibly yielding slightly more technical formulas). An extension to non-selfadjoint problems (including also nonlinear cost functions) could be realized taking into account Remark 6. An extension of models based on the Sherman–Morrison–Woodbury formula to other linear PDE constraints is straightforward since the structure of the discretized problem is the same as for our model problem. The topological derivative model can be extended to other PDE constraints following the general systematic procedure presented in [20].
- In this paper, we always assumed a structured mesh of a certain mesh topology to be given. While this is a common assumption made in many publications on topology optimization, an extension to general meshes with arbitrary element shapes and sizes would be an interesting topic of future research. In this setting, one might want to parametrize the shape of triangles. Then one could perform the precomputation

for a (small) number of sample triangle shapes and interpolate their data in order to treat a family of element shapes.

- The extension of the proposed approaches to nonlinear PDE constraints such as nonlinear elasticity or nonlinear magnetostatics is another interesting yet challenging task. Also, here, the general procedure for obtaining topological derivatives [20] could be used to establish a model similar to $\hat{\mathcal{J}}_{\mathrm{TDnum}}$.
- Finally, the ultimate goal of this research is to obtain good approximate subproblems in an iterative optimization algorithm. While solving the actual optimization problem was beyond the scope of this paper and subject of future research, we mention that this can be carried out in a similar way to [19]. In particular, in [19] it was shown that a sequential global programming approach with a diagonal approximation of a Sherman–Morrison–Woodbury model was superior to the well-established method of moving asymptotes (MMA) [17] in terms of both number of optimization iterations and quality of obtained solutions. A similar or even better behavior is expected when replacing the diagonal approximation model to our model $\hat{\mathcal{J}}_{\mathrm{SMWapprox}}$.

Acknowledgements This work has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1411 (project ID 416229255) and SFB 814 (project ID 61375930). The work of P. Gangl is supported by the joint DFG/FWF Collaborative Research Centre CREATOR (CRC – TRR361/F90) at TU Darmstadt, TU Graz and JKU Linz.

Funding Open access funding provided by Österreichische Akademie der Wissenschaften

Data availability The code and data used for all numerical experiments is available at https://zenodo.org/doi/10.5281/zenodo.11070440.

References

1. Delfour, M.C., Zolésio, J.-P. Shapes and Geometries, Metrics, analysis, differential calculus, and optimization. In: Advances in Design and Control, vol. 22. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, p. 622 (2011)
2. Sigmund, O., Maute, K.: Topology optimization approaches. Struct. Multidiscip. Optim. **48**, 1031–1055 (2013). https://doi.org/10.1007/s00158-013-0978-6
3. Allaire, G., Jouve, F., Toader, A.-M.: Structural optimization using sensitivity analysis and a level-set method. J. Comput. Phys. **194**(1), 363–393 (2004). https://doi.org/10.1016/j.jcp.2003.09.032
4. Amstutz, S., Andrä, H.: A new algorithm for topology optimization using a level-set method. J. Comput. Phys. **216**(2), 573–588 (2006)
5. Bendsøe, M.P., Sigmund, O.: Topology Optimization: Theory, Methods and Applications. Springer, Berlin (2003)
6. Allaire, G., Dapogny, C., Delgado, G., Michailidis, G.: Multi-phase structural optimization via a level set method. ESAIM: COCV **20**(2), 576–611 (2014). https://doi.org/10.1051/cocv/2013076

7. Gangl, P.: A multi-material topology optimization algorithm based on the topological derivative. Comput. Methods Appl. Mech. Eng. **366**, 113090 (2020). https://doi.org/10.1016/j.cma.2020.113090
8. Cherrière, T., Laurent, L., Hlioui, S., Louf, F., Duysinx, P., Geuzaine, C., Ahmed, H.B., Gabsi, M., Fernández, E.: Multi-material topology optimization using wachspress interpolations for designing a 3-phase electrical machine stator. Struct. Multidiscip. Optim. (2022). https://doi.org/10.1007/s00158-022-03460-1
9. Gangl, P., Gfrerer, M.H.: A unified approach to shape and topological sensitivity analysis of discretized optimal design problems. arXiv (2022). arxiv:2209.15491
10. Plotnikov, P.I., Sokolowski, J.: Geometric aspects of shape optimization. J. Geom. Anal. (2023). https://doi.org/10.1007/s12220-023-01252-7
11. Bourdin, B.: Filters in topology optimization. Int. J. Numer. Methods Eng. **50**(9), 2143–2158 (2001). https://doi.org/10.1002/nme.116
12. Semmler, J., Pflug, L., Stingl, M.: Material optimization in transverse electromagnetic scattering applications. SIAM J. Sci. Comput. **40**, 85–109 (2018). https://doi.org/10.1137/17M1127569
13. Fleury, C.: Structural weight optimization by dual methods of convex programming. Int. J. Numer. Methods Eng. **14**(12), 1761–1783 (1979). https://doi.org/10.1002/nme.1620141203
14. Bruyneel, M., Duysinx, P., Fleury, C.: A family of mma approximations for structural optimization. Struct. Multidiscip. Optim. **24**(4), 263–276 (2002). https://doi.org/10.1007/s00158-002-0238-7
15. Svanberg, K.: A class of globally convergent optimization methods based on conservative convex separable approximations. SIAM J. Optim. **12**(2), 555–573 (2002)
16. Fleury, C.: Conlin: An efficient dual optimizer based on convex approximation concepts. Struct. Optim. **1**(2), 81–89 (1989). https://doi.org/10.1007/BF01637664
17. Svanberg, K.: The method of moving asymptotes-a new method for structural optimization. Int. J. Numer. Methods Eng. **24**(2), 359–373 (1987). https://doi.org/10.1002/nme.1620240207
18. Novotny, A.A., Sokołowski, J.: Topological Derivatives in Shape Optimization. Springer, Berlin (2013)
19. Nees, N., Pflug, L., Mann, B., Stingl, M.: Multi-material design optimization of optical properties of particulate products by discrete dipole approximation and sequential global programming. Struct. Multidiscip. Optim. (2022). https://doi.org/10.1007/s00158-022-03376-w
20. Gangl, P., Sturm, K.: Automated computation of topological derivatives with application to nonlinear elasticity and reaction-diffusion problems. Comput. Methods Appl. Mech. Eng. **398**, 115288 (2022). https://doi.org/10.1016/j.cma.2022.115288
21. Deny, J., Lions, J.L.: Les espaces du type de Beppo Levi. Ann. Inst. Fourier, Grenoble **5**, 305–3701955 (1953–54)
22. Gangl, P., Sturm, K.: A simplified derivation technique of topological derivatives for quasi-linear transmission problems. ESAIM Control Optim. Calc. Var. **26**, 106–20 (2020). https://doi.org/10.1051/cocv/2020035
23. Ammari, H., Kang, H.: Polarization and Moment Tensors. Applied Mathematical Sciences. Springer, Berlin (2007)
24. Amstutz, S.: Sensitivity analysis with respect to a local perturbation of the material property. Asympt. Anal. **49**(1), 1–17 (2006)
25. Amstutz, S.: An introduction to the topological derivative. Eng. Comput. **39**(1), 3–33 (2021). https://doi.org/10.1108/ec-07-2021-0433
26. Amstutz, S., Gangl, P.: Topological derivative for the nonlinear magnetostatic problem. Electron. Trans. Numer. Anal. **51**, 169–218 (2019)
27. Gangl, P., Sturm, K.: Asymptotic analysis and topological derivative for 3D quasi-linear magnetostatics. ESAIM **55**, 853–875 (2021). https://doi.org/10.1051/m2an/2020060
28. Golub, G.H., Van Loan, C.F.: Matrix Computations. JHU Press, Baltimore (2013)
29. Amstutz, S., Dapogny, C., Ferrer, À.: A consistent relaxation of optimal design problems for coupling shape and topological derivatives. Numer. Math. **140**(1), 35–94 (2018)
30. Leugering, G., Nazarov, S., Schury, F., Stingl, M.: The Eshelby theorem and application to the optimization of an elastic patch. SIAM J. Appl. Math. **72**(2), 512–534 (2012). https://doi.org/10.1137/110823110
31. Schury, F., Greifenstein, J., Leugering, G., Stingl, M.: On the efficient solution of a patch problem with multiple elliptic inclusions. Optim. Eng. **16**(1), 225–246 (2015). https://doi.org/10.1007/s11081-014-9262-x
32. Nazarov, S.A.: Elasticity polarization tensor, surface enthalpy, and eshelby theorem. J. Math. Sci. **159**(2), 133–167 (2009). https://doi.org/10.1007/s10958-009-9432-0