Algorithmic aspects of multigrid methods for optimization in shape spaces

Martin Siebenborn[*]   Kathrin Welker[*]

Abstract

We examine the interaction of multigrid methods and shape optimization in appropriate shape spaces. Our aim is a scalable algorithm for application on supercomputers, which can only be achieved by mesh-independent convergence. The impact of discrete approximations of geometrical quantities, like the mean curvature, on a multigrid shape optimization algorithm with quasi-Newton updates is investigated. For the purpose of illustration, we consider a complex model for the identification of cellular structures in biology with minimal compliance in terms of elasticity and diffusion equations.

1 Introduction

PDE constrained shape optimization is becoming more and more suited for practical applications, e.g., acoustics [18], aerodynamics [15] and electrostatics [6]. A finite dimensional optimization problem can be obtained for example by representing shapes as splines. However, the connection of shape calculus with infinite dimensional spaces [5, 9, 22] leads to a more flexible approach. In recent work, it has been shown that PDE constrained shape optimization problems can be embedded in the framework of optimization on shape spaces. Finding a shape space and an associated metric is a challenging task and different approaches lead to various models. There exists no common shape space suitable for all applications. One possible approach is to define shapes as elements of a Riemannian manifold as proposed in [12, 13]. In [14], a survey of various suitable inner products is given, e.g., the curvature weighted Riemannian metric and the Sobolev metric. From a theoretical point of view this is attractive because algorithmic ideas from [1] can be combined with approaches from differential geometry. For example in [20], shape optimization is considered as optimization on a Riemannian shape manifold. This particular manifold contains only shapes with infinitely differentiable boundaries, which limits the practical applicability. From a computational point of view, usually finite element methods are used to discretize the PDE models where one has to deal with polygonal shape representations.

A well-established approach is to deal with shape derivatives in a so-called Hadamard form, i.e., in the form of integrals over the surface (cf. [15, 22]). An equivalent and intermediate result in the process of deriving Hadamard expressions is a volume expression of the shape derivative. One usually has to require additional regularity assumptions in order to transform volume into surface forms. In addition to saving analytical effort this makes volume expressions preferable over Hadamard forms, which is also utilized in [6]. In the case of the more attractive volume formulation, the shape manifold and the corresponding inner products mentioned above are not appropriate. One possible approach to use these formulations is given in

---

[*]Universität Trier, D-54286 Trier, Germany, Email: siebenborn@uni-trier.de, welker@uni-trier.de

[19]. Here an inner product, which is called Steklov-Poincaré metric, and a shape space are proposed. These are further considered in the following. The combination of this particular shape space and its associated inner product is an essential step towards applying efficient finite element solvers. This is especially important in order to avoid complicated load balancing for parallel computers with respect to both, the volume and surface elements. If there are many computationally costly operations on surfaces like the evaluation of the Sobolev metric, this is necessary. By using volume forms and Steklov-Poincaré-type metrics one only has to consider standard finite element assemblies and a classical load balancing with respect to volume elements.

In general, practical applications necessitate very fine discretizations. Here one can observe mesh-dependence at several points, e.g., the number of iterations of the PDE solver and the overall optimization loop. The same holds for discrete approximations of geometrical quantities, like the mean curvature (cf. [11]), which strongly depend on the chosen mesh. Especially a perimeter regularization is affected by the increasing values of the discretized mean curvature within a multgrid framework. In particular, scalable algorithms for application on supercomputers can only be achieved by mesh-independent convergence of both, the PDE simulation and the overall optimization loop. In order to come up with such an algorithm, one has to apply on the one hand multigrid preconditioner to the PDE solver and on the other hand quasi-Newton methods or something even more sophisticated for the optimization. In this paper, we examine the interaction of multigrid and shape optimization based on Steklov-Poincaré metrics and the corresponding shape space. We focus on this particular combination because it is well-suited for a large-scale finite element algorithm with quasi-Newton techniques.

This paper has the following structure. In section 2, besides a short review on the background of shape optimization, PDE models are formulated. Furthermore, the corresponding shape derivatives are given. Section 3 summarizes the entire process from shape derivatives to a complete optimization algorithm in suitable shape spaces. A numerical test framework and results are presented in section 4.

2   Model formulations and their shape derivatives

After setting up notation and terminology we formulate the model problems. They are motivated by finding the optimal design of cellular structures. In the third part of this section we deduce the shape derivatives of the model problems.

2.1   Notations and definitions

One focus in shape optimization is to investigate shape functionals. First, we define such a functional.

Definition 2.1 (Shape functional)**.** *Let D denote a non-empty subset of $\mathbb{R}^d$, where $d \in \mathbb{N}$. Moreover, $\mathcal{A} \subset \{\Omega \colon \Omega \subset D\}$ denotes a set of subsets. A function*

$$J \colon \mathcal{A} \to \mathbb{R}, \ \Omega \mapsto J(\Omega)$$

is called a shape functional.

In the following, let D be as in the definition above. Moreover, let $\{F_t\}_{t \in [0,T]}$ be a family of mappings $F_t \colon \overline{D} \to \mathbb{R}^d$ such that $F_0 = id$, where $T > 0$. This family transforms a domain $\Omega \subset D$ into *perturbed domains*

$$\Omega_t := F_t(\Omega) = \{F_t(x) \colon x \in \Omega\} \text{ with } \Omega_0 = \Omega$$

and the boundary $\Gamma = \partial\Omega$ into *perturbed boundaries*

$$\Gamma_t := F_t(\Gamma) = \{F_t(x)\colon x \in \Gamma\} \text{ with } \Gamma_0 = \Gamma.$$

Considering the domain Ω as a collection of material particles, which are changing their position in the time-interval $[0, T]$, the family $\{F_t\}_{t\in[0,T]}$ describes the motion of each particle. This means that at time $t \in [0, T]$ a particle $x \in \Omega$ has the new position $x_t := F_t(x) \in \Omega_t$ with $x_0 = x$. The motion of each such particle x can be described by the *velocity method* or by the *perturbation of identity*. Let V be a sufficiently smooth vector field. For V the velocity method defines the family of the above-mentioned mappings as the flow $F_t(x) := \xi(t, x)$, which is determined by the following initial value problem:

$$\frac{d\xi(t, x)}{dt} = V(\xi(t, x))$$
$$\xi(0, x) = x$$

The perturbation of identity is defined by $F_t(x) := x + tV(x)$ and used in the following.

This paper deals with PDE constrained shape optimization problems, i.e., shape optimization problems constrained by equations involving an unknown function of two or more variables and at least one partial derivative of this function. Such a problem is given by

$$\min_\Omega J(\Omega),$$

where J is a shape functional additionally depends on a solution of a PDE. To solve these problems, we need their shape derivatives:

Definition 2.2 (Shape derivative)**.** *Let $D \subset \mathbb{R}^d$ be open, where $d \geq 2$ is a natural number. Moreover, let $k \in \mathbb{N} \cup \{\infty\}$ and let $\Omega \subset D$ be measurable. The Eulerian derivative of a shape functional J at Ω in direction $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$ is defined by*

$$DJ(\Omega)[V] := \lim_{t\to 0^+} \frac{J(\Omega_t) - J(\Omega)}{t}. \tag{2.1}$$

If for all directions $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$ the Eulerian derivative (2.1) exists and the mapping

$$G(\Omega)\colon \mathcal{C}_0^k(D, \mathbb{R}^d) \to \mathbb{R}, \ V \mapsto DJ(\Omega)[V]$$

is linear and continuous, the expression $DJ(\Omega)[V]$ is called the shape derivative of J at Ω in direction $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$. In this case, J is called shape differentiable of class $\mathcal{C}^k$ at Ω.

For a detailed introduction into shape calculus, we refer to the monographs [5, 22]. In particular, [22] states that shape derivatives can always be expressed as boundary integrals due to the Hadamard structure theorem [22, theorem 2.27]. In many cases, the shape derivative arises in two equivalent notational forms:

$$DJ_\Omega[V] := \int_\Omega F(x)V(x)\,dx \qquad \text{(volume formulation)} \qquad (2.2)$$

$$DJ_\Gamma[V] := \int_\Gamma f(s)V(s)^T n(s)\,ds \qquad \text{(surface formulation)} \qquad (2.3)$$

Here F is a (differential) operator acting linearly on the vector field V and $f \in L^1(\Gamma)$ with $DJ_\Omega[V] = DJ(\Omega)[V] = DJ_\Gamma[V]$.

In general, we have to deal with so-called material derivatives in order to derive shape derivatives. The definition of these shape derivatives is given in the following. For a material derivative free approach we refer to [23].

3

Definition 2.3 (Material derivative)**.** *Let* Ω, Ω_t, F_t *and* T *be as above. Moreover, let* $\{p_t \colon \Omega_t \to \mathbb{R} \colon t \leq T\}$ *denote a family of mappings. The material derivative of a generic function* $p(= p_0) \colon \Omega \to \mathbb{R}$ *at* $x \in \Omega$ *is denoted by* $D_m p$ *or* $\dot{p}$ *and given by the derivative of the composed function* $p_t \circ F_t \colon \Omega \to \Omega_t \to \mathbb{R}$ *defined in the fixed domain* Ω*, i.e.,*

$$\dot{p}(x) := \lim_{t \to 0^+} \frac{(p_t \circ F_t)(x) - p(x)}{t} = \frac{d^+}{dt} (p_t \circ F_t)(x) \Big|_{t=0}.$$

In the following, let $p \colon \Omega \to \mathbb{R}$ be a function. The classical chain rule for differentiation applied to $\dot{p}$ gives the relation between material and shape derivatives. The shape derivative of p in the direction of a vector field V is denoted by p' and given by

$$p' = \dot{p} - V^T \nabla p \quad \text{in } \Omega. \tag{2.4}$$

In subsection 2.3, the following rules for the material derivative are needed to derive shape derivatives of objective functions depending on solutions of PDEs. For the material derivative the product rule holds, i.e.,

$$D_m(p\,q) = D_m p\,q + p\,D_m q. \tag{2.5}$$

While the shape derivative commutes with the gradient, the material derivative does not, but the following equality, which is proven in [2], holds:

$$D_m \nabla p = \nabla D_m p - \nabla V^T \nabla p. \tag{2.6}$$

The concept of material and shape derivatives of a function $p \colon \Omega \to \mathbb{R}$ can be extended to its boundary $\Gamma = \partial \Omega$. We mention only a few aspects. For more details we refer to the literature, e.g., [17]. Let $z \colon \Gamma \to \mathbb{R}$ be the trace on the boundary Γ of p. In this setting, the boundary shape derivative z' is defined by

$$z' = \dot{p} - V^T \nabla_\Gamma p, \tag{2.7}$$

where ∇_Γ denotes the tangential gradient given by

$$\nabla_\Gamma p = \nabla p - \frac{\partial p}{\partial n} n.$$

Here $\frac{\partial}{\partial n}$ denotes the derivative normal to Γ. Combining (2.4) with (2.7) gives the correlation of boundary and domain shape derivatives:

$$z' = p' + V^T \frac{\partial p}{\partial n} n$$

In order to deduce shape derivative formulas we have to consider the derivative of perturbed objective functions:

$$\frac{d^+}{dt} \left(\int_{\Omega_t} p_t \, dx_t \right) \Big|_{t=0} \qquad \text{and} \qquad \frac{d^+}{dt} \left(\int_{\Gamma_t} z_t \, ds_t \right) \Big|_{t=0},$$

where $\Omega, \Omega_t, p, p_t, \Gamma_t$ are as above and $z_t \colon \Gamma_t \to \mathbb{R}$ denotes a mapping. They are given by

$$\frac{d^+}{dt} \left(\int_{\Omega_t} p_t \, dx_t \right) \Big|_{t=0} = \int_\Omega \dot{p} + \mathrm{div}(V)p \, dx, \tag{2.8}$$

$$\frac{d^+}{dt} \left(\int_{\Gamma_t} z_t \, ds_t \right) \Big|_{t=0} = \int_\Gamma \dot{z} + \mathrm{div}_\Gamma(V)z \, ds, \tag{2.9}$$

where

$$\mathrm{div}_\Gamma(V) := \mathrm{div}(V) - n^T \frac{\partial V}{\partial n}$$

denotes the tangential divergence of V. For their proofs we refer to the literature, e.g., [8, 9, 28].

2.2 Problem formulations

Let $\Omega \subset D$ be a bounded Lipschitz domain with boundary $\partial\Omega$. This domain is assumed to be a cuboid and partitioned in a subdomain $\Omega_{\text{out}} \subset \Omega$ and a finite number of disjoint subdomains $\Omega_i \subset \Omega$ with boundaries $\Gamma_i := \partial\Omega_i$ such that $\Omega_{\text{out}} \bigcup (\cup_{i\in\mathbb{N}}\Omega_i) \bigcup (\cup_{i\in\mathbb{N}}\Gamma_i) = \Omega$ and $\Gamma_{\text{bottom}} \uplus \Gamma_{\text{sides}} \uplus \Gamma_{\text{top}} = \partial\Omega$ ($=: \Gamma_{\text{out}}$), where $\uplus$ denotes the disjoint union. The union of all domains Ω_i is denoted by $\Omega_{\text{int}} := \bigcup_{i\in\mathbb{N}} \Omega_i$ and the union of all boundaries Γ_i is called the interface and denoted by $\Gamma_{\text{int}} := \bigcup_{i\in\mathbb{N}} \Gamma_i$. The outer normal vector field to Ω_{int} is given by n. Figure 1 illustrates this situation. In our setting, Ω is meant to be composed of two distinct materials, one in Ω_{int} and one in Ω_{out}.

Let $\nu_l \in \mathbb{R}$ be arbitrary constants, where $l \in \{1, \cdots, 4\}$. For the objective function

$$J(y, u, \Omega) = j_1(u, \Omega) + j_2(y, \Omega) + j_3(\Omega) + j_4(\Omega) \tag{2.10}$$

with

$$j_1(u, \Omega) := \nu_1 \int_\Omega \sigma(u) : \epsilon(u) \, dx,$$

$$j_2(y, \Omega) := \frac{\nu_2}{2} \int_0^T \int_\Omega (y - \bar{y})^2 \, dx \, dt,$$

$$j_3(\Omega) := \nu_3 \int_{\Omega_{\text{out}}} 1 \, dx,$$

$$j_4(\Omega) := \nu_4 \int_{\Gamma_{\text{int}}} 1 \, ds$$

we consider the following PDE constrained optimization problem in strong form:

$$\min_{\Gamma_{\text{int}}} J(y, u, \Omega) \tag{2.11}$$

subject to the following constraints:

$$-\operatorname{div}\sigma(u) = 0 \quad \text{in } \Omega \tag{2.12}$$

$$u = 0 \quad \text{on } \Gamma_{\text{bottom}} \tag{2.13}$$

$$\sigma(u) \cdot n = f \quad \text{on } \Gamma_{\text{top}} \cup \Gamma_{\text{sides}} \tag{2.14}$$

$$\frac{\partial y}{\partial t} - \operatorname{div}(k\nabla y) = 0 \quad \text{in } \Omega \times (0, T] \tag{2.15}$$

$$y = 1 \quad \text{on } \Gamma_{\text{top}} \times (0, T] \tag{2.16}$$

$$\frac{\partial y}{\partial n} = 0 \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{sides}}) \times (0, T] \tag{2.17}$$

$$y = 0 \quad \text{in } \Omega \times \{0\} \tag{2.18}$$

Equations (2.12)-(2.14) describe the linear elasticity model, where $\sigma := \lambda \operatorname{Tr}(\epsilon)I + 2\mu\epsilon$ denotes the stress tensor and $\epsilon = \frac{1}{2}\left(\nabla u + \nabla u^T\right)$ the strain tensor with respect to the Lamé parameters λ and μ. We further assume

$$f = \begin{cases} 0 & \text{on } \Gamma_{\text{sides}} \\ f_{\text{top}} & \text{on } \Gamma_{\text{top}} \end{cases}, \tag{2.19}$$

where f_{top} is a force at the top boundary of the domain, which leads to the deformation u. A diffusion model is given by (2.15)-(2.18) with a jumping permeability coefficient k. The different properties of the two materials with respect to elasticity

Figure 1: Illustration of the domain Ω

and permeability are modelled by the following choice of coefficients:

$$k := \begin{cases} k_{\text{out}} & \text{in } \Omega_{\text{out}} \\ k_{\text{int}} & \text{in } \Omega_{\text{int}} \end{cases}, \quad \lambda := \begin{cases} \lambda_{\text{out}} & \text{in } \Omega_{\text{out}} \\ \lambda_{\text{int}} & \text{in } \Omega_{\text{int}} \end{cases}, \quad \mu := \begin{cases} \mu_{\text{out}} & \text{in } \Omega_{\text{out}} \\ \mu_{\text{int}} & \text{in } \Omega_{\text{int}} \end{cases}.$$

Due to the jump in these coefficients, the formulations (2.12) and (2.15) are to be understood only formally. The objective function J is composed of the four shape functionals j_l. Here j_1 corresponds to the minimization of the compliance of the composite material in the domain Ω. With j_2 the diffusion model is fitted to data measurements $\bar{y}$. We assume for the observation $\bar{y} \in L^2\left(0, T; L^2(\Omega)\right)$. Since we are interested in a space-filling design of the inclusions Ω_{int}, the purpose of the functional j_3 is to minimize the volume of Ω_{out}. This is equivalent to maximizing Ω_{int} if the outer shape Γ_{out} is fixed. j_4 is a perimeter regularization. Of course, the integral in j_4 has to be understood as the sum of the integrals over Γ_{int}. We formulate explicitly the continuity of the state and of the flux at the interface as

$$\llbracket u \rrbracket = 0, \qquad \llbracket \sigma(u) \cdot n \rrbracket = 0 \qquad \text{on } \Gamma_{\text{int}}, \tag{2.20}$$

$$\llbracket y \rrbracket = 0, \qquad \left\llbracket k \frac{\partial y}{\partial n} \right\rrbracket = 0 \qquad \text{on } \Gamma_{\text{int}} \times (0, T], \tag{2.21}$$

where the jump symbol $\llbracket \cdot \rrbracket$ denotes the discontinuity across the interface Γ_{int} and is defined by $\llbracket v \rrbracket := v\big|_{\Omega_{\text{int}}} - v\big|_{\Omega_{\text{out}}}$ for $v \in \Omega$.

Remark 2.4. *Equations (2.12)-(2.14) admit a solution $u \in H^1(\Omega, \mathbb{R}^d)$ and (2.15)-(2.18) have a solution $y \in L^2\left(0, T; H^1(\Omega)\right)$. However, it turns out that the restrictions of these solutions to Ω_{int} and Ω_{out} have a higher regularity (cf. for example [9, proposition 11.7]). To be more precise, $u\big|_{\Omega_{int}} \in H^2(\Omega_{int})$, $u\big|_{\Omega_{out}} \in H^2(\Omega_{out}, \mathbb{R}^d)$, $y\big|_{\Omega_{int}} \in L^2\left(0, T; H^2(\Omega_{int})\right)$ and $y\big|_{\Omega_{out}} \in L^2\left(0, T; H^2(\Omega_{out})\right)$. In this setting, the integral over Ω has to be understood as the sum of the integrals over Ω_{int} and Ω_{out}. In the following, the integral over Ω always denotes this sum. Thus, it is guaranteed that $\frac{\partial u|_{\Omega_{int}}}{\partial n} \in H^{1/2}\left(\Gamma_{\text{int}}, \mathbb{R}^d\right)$ and $\frac{\partial y|_{\Omega_{int}}}{\partial n} \in L^2\left(0, T; H^{1/2}(\Gamma_{\text{int}})\right)$ by the trace theorem for Sobolev spaces.*

The boundary value problem (2.12)-(2.14) is given in its weak form by

$$a_1(u, w) = b_1(u, w_1, w_2), \ \forall w \in H^1\left(\Omega, \mathbb{R}^d\right) \tag{2.22}$$

and for all $w_1 \in H^{-1/2}\left(\Gamma_{\text{bottom}}, \mathbb{R}^d\right)$, $w_2 \in H^{1/2}\left(\Gamma_{\text{sides}} \cup \Gamma_{\text{top}}, \mathbb{R}^d\right)$. Here the bilinear form $a_1(u, w)$ is given by

$$
\begin{aligned}
a_1(u, w) := \int_\Omega \sigma(u) : \epsilon(w) \, dx - \int_{\Gamma_{\text{int}}} \left[\!\left[(\sigma(u) \cdot n)^T w\right]\!\right] dx \\
- \int_{\Gamma_{\text{out}}} (\sigma(u) \cdot n)^T w \, ds
\end{aligned}
\tag{2.23}
$$

and $b_1(u, w_1, w_2)$ is defined by

$$b_1(u, w_1, w_2) := \int_{\Gamma_{\text{bottom}}} w_1^T u \, ds + \int_{\Gamma_{\text{out}} \backslash \Gamma_{\text{bottom}}} w_2^T (\sigma(u) \cdot n - f) \, ds. \tag{2.24}$$

Moreover, problem (2.15)-(2.18) is written in weak form as

$$a_2(y, z) = b_2(y, z_1, z_2), \ \forall z \in W\left(0, T; H^1(\Omega)\right) \tag{2.25}$$

and for all $z_1 \in L^2\left(0, T; H^{-1/2}(\Gamma_{\text{top}})\right)$, $z_2 \in L^2\left(0, T; H^{1/2}(\Gamma_{\text{bottom}} \cup \Gamma_{\text{sides}})\right)$ as in [20]. Here the bilinear form $a_2(y, z)$ is given by

$$
\begin{aligned}
a_2(y, z) := \int_\Omega y(T, x) \, z(T, x) \, dx - \int_0^T \int_\Omega \frac{\partial z}{\partial t} y + k \nabla y^T \nabla z \, dx \, dt \\
- \int_0^T \int_{\Gamma_{\text{int}}} \left[\!\left[k \frac{\partial y}{\partial n} z\right]\!\right] ds \, dt - \int_0^T \int_{\Gamma_{\text{out}}} k_{\text{out}} \frac{\partial y}{\partial n} z \, ds \, dt
\end{aligned}
\tag{2.26}
$$

and $b_2(y, z_1, z_2)$ by

$$b_2(y, z_1, z_2) := \int_0^T \int_{\Gamma_{\text{top}}} z_1(y - 1) \, ds \, dt + \int_0^T \int_{\Gamma_{\text{out}} \backslash \Gamma_{\text{top}}} z_2 \frac{\partial y}{\partial n} \, ds \, dt. \tag{2.27}$$

For properties of the function spaces, we refer to the literature, e.g., [7, 26].

The Lagrangian of (2.10)-(2.18) is defined by

$$\mathscr{L}(y, u, w, z, \Omega) = \mathscr{L}_1(u, w, \Omega) + \mathscr{L}_2(y, z, \Omega) + \mathscr{L}_3(\Omega) + \mathscr{L}_4(\Omega) \tag{2.28}$$

with

$$
\begin{aligned}
\mathscr{L}_1(u, w, \Omega) &:= j_1(u, \Omega) + a_1(u, w) - b_1(u, w_1, w_2), \\
\mathscr{L}_2(y, z, \Omega) &:= j_2(y, \Omega) + a_2(y, z) - b_2(y, z_1, z_2), \\
\mathscr{L}_3(\Omega) &:= j_3(\Omega), \\
\mathscr{L}_4(\Omega) &:= j_4(\Omega).
\end{aligned}
$$

The adjoint problem to (2.12)-(2.14), which we obtain from differentiating the Lagrangian $\mathscr{L}_1$ with respect to u, is given in strong form by

$$-\operatorname{div} \sigma(w) = 0 \quad \text{in } \Omega \tag{2.29}$$

$$w = 0 \quad \text{on } \Gamma_{\text{bottom}} \tag{2.30}$$

$$\sigma(w) \cdot n = -\nu_1 f \quad \text{on } \Gamma_{\text{sides}} \cup \Gamma_{\text{top}} \tag{2.31}$$

$$w_1 = \sigma(w) \cdot n \quad \text{on } \Gamma_{\text{bottom}} \tag{2.32}$$

$$w_2 = -w \quad \text{on } \Gamma_{\text{sides}} \cup \Gamma_{\text{top}} \tag{2.33}$$

and the state equation to (2.12)-(2.14), which we get by differentiating the Lagrangian $\mathscr{L}_1$ with respect to w, is given in strong form by

$$- \operatorname{div} \sigma(u) = 0 \quad \text{in } \Omega. \tag{2.34}$$

Moreover, the adjoint problem to (2.15)-(2.18), which we obtain from differentiating the Lagrangian $\mathscr{L}_2$ with respect to y, is given in strong form by

$$-\frac{\partial z}{\partial t} - \operatorname{div}(k\nabla z) = -\nu_2(y - \overline{y}) \quad \text{in } \Omega \times [0, T] \tag{2.35}$$

$$\frac{\partial z}{\partial n} = 0 \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{sides}}) \times [0, T] \tag{2.36}$$

$$z = 0 \quad \text{on } \Gamma_{\text{top}} \times [0, T] \tag{2.37}$$

$$z = 0 \quad \text{in } \Omega \times \{T\} \tag{2.38}$$

$$z_1 = k_{\text{out}} \frac{\partial z}{\partial n} \quad \text{on } \Gamma_{\text{top}} \times [0, T] \tag{2.39}$$

$$z_2 = -k_{\text{out}} z \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{sides}}) \times [0, T] \tag{2.40}$$

and the state equation to (2.15)-(2.18), which we get by differentiating the Lagrangian $\mathscr{L}_2$ with respect to z, is given in strong form by

$$\frac{\partial y}{\partial t} - \operatorname{div}(k\nabla y) = 0 \quad \text{in } \Omega \times (0, T]. \tag{2.41}$$

We formulate explicitly the interface condition of (2.29)-(2.33) and (2.35)-(2.40) by

$$\llbracket w \rrbracket = 0, \qquad \llbracket \sigma(w) \cdot n \rrbracket = 0 \qquad \text{on } \Gamma_{\text{int}}, \tag{2.42}$$

$$\llbracket z \rrbracket = 0, \qquad \left\llbracket k\frac{\partial z}{\partial n} \right\rrbracket = 0 \qquad \text{on } \Gamma_{\text{int}} \times [0, T]. \tag{2.43}$$

There are a lot of options to prove shape differentiability of shape functionals, which depend on a solution of a PDE, and to derive the shape derivative of a PDE constrained shape optimization problem. The min-max approach [5], the chain rule approach [22], the Lagrange method of Céa [3] and the rearrangement method [10] have to be mentioned in this context. A nice overview about these approaches is given in [24]. The next subsection deals with the shape derivatives of the models above.

2.3 Shape derivatives

In the sequel, we deduce the shape derivative of each shape functional j_l, where $l \in \{1, 2, 3, 4\}$. The sum of these four shape derivatives is the shape derivative of J. Note that we need only the volume form of the shape derivative of j_l with $l \in \{1, 2, 3\}$ with respect to the Steklov-Poincaré metric and the optimization techniques established in [19, 28]. The existence of these shape derivatives is given by the theorem of Correa and Seeger [4, theorem 2.1].

The following theorem gives a representation of the shape derivative of j_1 expressed as volume integral.

Theorem 2.5. *Let $u \in H^1(\Omega, \mathbb{R}^d)$ denote the weak solution of (2.12)-(2.14). Moreover, $w \in H^1(\Omega, \mathbb{R}^d)$ denote the weak solution of the adjoint equation (2.29)-(2.33). Then the shape derivative of j_1 in direction V is given by*

$$\boxed{\begin{aligned} Dj_1(u, \Omega)[V] = \int_\Omega &-\sigma(u) : \left(\nabla V^T \nabla w\right) - \sigma(w) : \left(\nabla V^T \nabla u\right) \\ &+ \operatorname{div}(V)\left(\nu_1 \sigma(u) : \epsilon(u) + \sigma(w) : \nabla u\right) \, dx. \end{aligned}} \tag{2.44}$$

Proof. We consider the Lagrangian $\mathscr{L}_1$. In analogy to [5, chapter 10, subsection 5.2], we can verify that

$$j_1(u, \Omega) = \min_{u \in H^1(\Omega, \mathbb{R}^d)} \max_{w \in H^1(\Omega, \mathbb{R}^d)} \mathscr{L}_1(u, w, \Omega) \tag{2.45}$$

holds. We apply the theorem of Correa and Seeger on the right-hand side of (2.45). The verification of the assumptions of this theorem can be checked in the same way as in [5, chapter 10, subsection 6.4].

Applying the rules for differentiating volume and surface integrals given in (2.8) and (2.9) yields

$$
\begin{aligned}
& D\mathscr{L}_1(u, w, \Omega)[V] \\
&= \int_\Omega \nu_1 D_m \left(\sigma(u) : \epsilon(u) \right) + D_m \left(\sigma(u) : \epsilon(w) \right) \\
&\qquad + \operatorname{div}(V) \left(\nu_1 \sigma(u) : \epsilon(u) + \sigma(u) : \epsilon(w) \right) dx \\
&\quad - \int_{\Gamma_{\text{int}}} D_m \left(\left[\!\left[(\sigma(u) \cdot n)^T w \right]\!\right] \right) + \operatorname{div}_{\Gamma_{\text{int}}}(V) \left[\!\left[(\sigma(u) \cdot n)^T w \right]\!\right] ds \\
&\quad - \int_{\Gamma_{\text{out}}} D_m \left((\sigma(u) \cdot n)^T w \right) + \operatorname{div}_{\Gamma_{\text{out}}}(V) \left((\sigma(u) \cdot n)^T w \right) ds \\
&\quad - \int_{\Gamma_{\text{bottom}}} D_m \left(w_1^T u \right) + \operatorname{div}_{\Gamma_{\text{bottom}}}(V) w_1^T u \, ds \\
&\quad - \int_{\Gamma_{\text{out}} \backslash \Gamma_{\text{bottom}}} D_m \left(w_2^T \left(\sigma(u) \cdot n - f \right) \right) \\
&\qquad\qquad + \operatorname{div}_{\Gamma_{\text{out}} \backslash \Gamma_{\text{bottom}}}(V) \, w_2^T \left(\sigma(u) \cdot n - f \right) ds.
\end{aligned}
\tag{2.46}
$$

Note that

$$\sigma(u) : \epsilon(w) = \sigma(u) : \nabla w \tag{2.47}$$

holds because $\sigma(u)$ is symmetric (cf. [21]). Moreover, we have

$$D_m(\sigma(u) : \nabla w) = \mu D_m \left((\nabla u + \nabla u^T) : \nabla w \right) + \lambda D_m \left(\operatorname{div}(u) \operatorname{div}(w) \right) \tag{2.48}$$

due to the definition of σ and ϵ. Combining (2.5) with (2.6) in $\mathbb{R}^d$ yields

$$D_m(\operatorname{div}(u)) = D_m(I : \nabla u) = \operatorname{div}(\dot{u}) - I : (\nabla V^T \nabla u). \tag{2.49}$$

By applying the product rule and appropriate transformations we get

$$
\begin{aligned}
& D_m(\sigma(u) : \epsilon(w)) \\
&= \sigma(\dot{u}) : \nabla w + \sigma(u) : \nabla \dot{w} - \sigma(u) : (\nabla V^T \nabla w) - \sigma(w) : (\nabla V^T \nabla u)
\end{aligned}
\tag{2.50}
$$

due to (2.6) and (2.47)-(2.49). Since the outer boundary Γ_{out} is fixed, we can choose the deformation vector field V equals zero in small neighbourhoods of Γ_{out}. Moreover, each material derivative in small neighbourhoods of Γ_{out} is equal to zero. Thus, the outer integrals in (2.51) are not further considered. From (2.46) we obtain the following under consideration of the adjoint and design equation in weak form and by appropriate transformations:

$$
\begin{aligned}
D\mathscr{L}_1(u, w, \Omega)[V] &= \int_\Omega -\sigma(u) : \left(\nabla V^T \nabla w \right) - \sigma(w) : \left(\nabla V^T \nabla u \right) \\
&\qquad + \operatorname{div}(V) \left(\nu_1 \sigma(u) : \epsilon(u) + \sigma(w) : \nabla u \right) dx \\
&\quad + \int_{\Gamma_{\text{int}}} \left[\!\left[(\sigma(w) \cdot n)^T \dot{u} - D_m \left((\sigma(u) \cdot n)^T \right) w \right]\!\right] \\
&\qquad + \operatorname{div}_{\Gamma_{\text{int}}}(V) \left[\!\left[(\sigma(u) \cdot n)^T w \right]\!\right] ds
\end{aligned}
\tag{2.51}
$$

9

Due to the identity $[\![\Phi\Psi]\!] = [\![\Phi]\!]\,\Psi_{\text{out}} + \Phi_{\text{int}}\,[\![\Psi]\!] = \Phi_{\text{out}}\,[\![\Psi]\!] + [\![\Phi]\!]\,\Psi_{\text{int}}$, which implies

$$[\![\Phi\Psi]\!] = 0 \ \text{if} \ [\![\Phi]\!] = 0 \wedge [\![\Psi]\!] = 0,$$

and the interface conditions (2.20) and (2.42), the interface integral vanishes in (2.51). By applying the theorem of Correa and Seeger, we obtain (2.44). $\qquad\square$

Now, we consider the shape functional j_2 and the diffusion problem (2.15)-(2.18). If we consider $\mathscr{L}_2$, its shape derivative can be deduced analogously to the computations in [20]. Thus, we get

$$
\boxed{
\begin{aligned}
Dj_2(y,\Omega)[V] = \int_0^T \int_\Omega &-k\nabla y^T\left(\nabla V + \nabla V^T\right)\nabla z \\
&+ \operatorname{div}(V)\left(\frac{\nu_2}{2}(y-\bar{y})^2 + \frac{\partial y}{\partial t}z + k\nabla y^T\nabla z\right)\,dx\,dt.
\end{aligned}
}
\tag{2.52}
$$

The shape derivative of j_3, which is responsible for a space-filling design of Ω_{int}, is given by

$$
\boxed{Dj_3(\Omega)[V] = \nu_3 \int_{\Omega_{\text{out}}} \operatorname{div}(V)\,dx.}
\tag{2.53}
$$

It results directly from an application of (2.8).

Moreover, the objective function J includes the perimeter regularization term j_4. We get the shape derivative of this regularization term by applying (2.9):

$$
\begin{aligned}
Dj_4(\Omega)[V] &= \nu_4 \int_{\Gamma_{\text{int}}} \operatorname{div}_{\Gamma_{\text{int}}}(V)\,ds = \nu_4 \int_{\Gamma_{\text{int}}} \operatorname{div}_{\Gamma_{\text{int}}}(\langle V,n\rangle\, n)\,ds \\
&= \nu_4 \int_{\Gamma_{\text{int}}} \langle V,n\rangle\, \operatorname{div}_{\Gamma_{\text{int}}}(n)\,ds,
\end{aligned}
$$

i.e.,

$$
\boxed{Dj_4(\Omega)[V] = \nu_4 \int_{\Gamma_{\text{int}}} \kappa\, \langle V,n\rangle\,ds,}
\tag{2.54}
$$

where $\kappa := \operatorname{div}_{\Gamma_{\text{int}}}(n)$ denotes the mean curvature of Γ_{int}.

So far, we have deduced derivative. However, in order to optimize on shape spaces, we need the gradient with respect to the inner product under consideration. In the next section, we comment on shape spaces and metrics.

3 Optimization based on Steklov-Poincaré metrics

As pointed out for example in [19, 20, 28], PDE constrained shape optimization problems can be embedded in the framework of optimization on shape spaces. This section summarizes the way from shape derivatives to an entire optimization algorithm in a suitable shape space.

In our setting, we think of shapes as boundary contours of deforming objects. We consider so-called prior shapes Γ_0. These are boundaries $\Gamma_0 = \partial\mathcal{X}_0$ of connected and compact subsets $\mathcal{X}_0 \subset \Omega \subset \mathbb{R}^d$ with $\mathcal{X}_0 \neq \emptyset$, where Ω denote bounded Lipschitz domains. Let $\mathcal{X}_0$ be Lipschitz domains and called prior sets. Shapes – in our setting – arise from H^1-deformations of such prior sets $\mathcal{X}_0$. These H^1-deformations, evaluated at a prior shape $\Gamma_0 = \partial\mathcal{X}_0$, give deformed shapes Γ if the deformations are injective and continuous. These shapes are called of class $H^{1/2}$ and proposed firstly in [19]. They are defined by

$$
\mathcal{B}^{1/2}(\Gamma_0,\mathbb{R}^d) := \mathcal{H}^{1/2}(\Gamma_0,\mathbb{R}^d)\big/\operatorname{Homeo}^{1/2}(\Gamma_0),
\tag{3.1}
$$

where $\mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d)$ is given by

$$
\mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d) := \{w \colon \Gamma_0 \to \Omega \colon \exists W \in H^1(\Omega, \Omega) \text{ s.t.}
$$
$$
W|_{\Gamma_0} \text{ injective, continuous, } W|_{\Gamma_0} = w\} \tag{3.2}
$$

and $\mathrm{Homeo}^{1/2}(\Gamma_0)$ is defined by

$$
\mathrm{Homeo}^{1/2}(\Gamma_0) := \{w \colon w \in \mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d), \, w \colon \Gamma_0 \to \Gamma_0 \text{ homeomorphism}\}. \tag{3.3}
$$

In the following, we assume that $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ has a manifold structure. If necessary, we can refine the space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ as already described in [28]. However, this conceivable limitation leaves the following theory untouched. If $\Gamma \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ is smooth enough to admit a normal vector field n, the following isomorphisms arise from definition (3.2):

$$
T_\Gamma \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d) \cong \{h \colon h = \phi n \text{ a.e., } \phi \in H^{1/2}(\Gamma) \text{ continuous}\}
$$
$$
\cong \{\phi \colon \phi \in H^{1/2}(\Gamma) \text{ continuous}\} \tag{3.4}
$$

We consider the following scalar products, the so-called *Steklov-Poincaré metrics* (cf. [19]):

$$
g^S \colon H^{1/2}(\Gamma_{\mathrm{int}}) \times H^{1/2}(\Gamma_{\mathrm{int}}) \to \mathbb{R},
$$
$$
(\alpha, \beta) \mapsto \langle \alpha, (S^{pr})^{-1} \beta \rangle = \int_{\Gamma_{\mathrm{int}}} \alpha(s) \cdot [(S^{pr})^{-1} \beta](s) \, ds \tag{3.5}
$$

Here S^{pr} denotes the projected Poincaré-Steklov operator which is given by

$$
S^{pr} \colon H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma), \, \alpha \mapsto (\gamma_0 U)^T n, \tag{3.6}
$$

where $\gamma_0 \colon H_0^1(\Omega, \mathbb{R}^d) \to H^{1/2}(\Gamma, \mathbb{R}^d), \, U \mapsto U|_\Gamma$ and $U \in H_0^1(\Omega, \mathbb{R}^d)$ solves the Neumann problem

$$
a(U, V) = \int_\Gamma \alpha \cdot (\gamma_0 V)^T n \, ds \quad \forall V \in H_0^1(\Omega, \mathbb{R}^d) \tag{3.7}
$$

with a being a symmetric and coercive bilinearform.

We are now able to formulate an optimization algorithm on the shape space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S. Before we can do that, we have to state its connection to shape calculus. Due to the Hadamard structure theorem (cf. [22, theorem 2.27]), there exists a scalar distribution r on the boundary of the domain under consideration. If we assume $r \in L^1(\Gamma)$ and $V|_\Gamma = \alpha n$, the shape derivative can be expressed as the boundary integral

$$
DJ_\Gamma[V] = \int_\Gamma \alpha \, r \, ds. \tag{3.8}
$$

Due to the isomorphism (3.4) and the expression (3.8), we can state the connection of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to the Steklov-Poincaré metric g^S to shape calculus. The distribution r is often called the *shape gradient*. This is confusing, because one has to note that gradients always depend on chosen scalar products defined on the space under consideration. It rather means that r is the usual L^2-shape gradient. However, we have to find a representation of the shape gradient with respect to g^S. Such a representation $h \in T_\Gamma \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d) \cong \{h \colon h \in H^{1/2}(\Gamma) \text{ continuous}\}$ is determined by

$$
g^S(\phi, h) = (r, \phi)_{L^2(\Gamma)}, \tag{3.9}
$$

Figure 2: Optimization algorithm

which is equivalent to

$$\int_\Gamma \phi(s) \cdot [(S^{pr})^{-1}h](s) \; ds = \int_\Gamma r(s)\phi(s) \; ds \tag{3.10}$$

for all continuous $\phi \in H^{1/2}(\Gamma)$. Based on the connection (3.9) we can formulate an algorithm to solve PDE constrained shape optimization problems on $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S. From (3.10) we get $h = S^{pr}r = (\gamma_0 U)^T n$, where $U \in H^1_0(\Omega, \mathbb{R}^d)$ solves

$$a(U, V) = \int_\Gamma r \cdot (\gamma_0 V)^T n \; ds = DJ_\Gamma[V] = DJ_\Omega[V] \quad \forall V \in H^1_0(\Omega, \mathbb{R}^d). \tag{3.11}$$

This means that we get the gradient representation h and the mesh deformation U all at once and that we have to solve

$$a(U, V) = b(V) \tag{3.12}$$

for all test functions V in the optimization algorithm, where b is a linear form and given by

$$b(V) := DJ_{\text{vol}}(\Omega)[V] + DJ_{\text{surf}}(\Omega)[V].$$

$J_{\text{surf}}(\Omega)$ denotes parts of the objective function leading to surface shape derivative expressions, e.g., perimeter regularizations. The shape derivative $DJ_{\text{surf}}(\Omega)[V]$ of these terms are incorporated as Neumann boundary conditions. Parts of the objective function leading to volume shape derivative expressions are denoted by $J_{\text{vol}}(\Omega)$. However, note that from a theoretical point of view the volume and surface shape derivative formulations have to be equal to each other for all test functions. Thus, $DJ_{\text{vol}}[V]$ is assembled only for test functions V whose support includes Γ, i.e.,

$$DJ_{\text{vol}}(\Omega)[V] = 0 \quad \forall V \text{ with } \text{supp}(V) \cap \Gamma = \emptyset.$$

We call (3.12) the *deformation equation*. The entire optimization algorithm is given in figure 2 and explained in detail in the next section.

In the setting of section 2, each Γ_i is a shape in the above sense, i.e., an element of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$. The volume formulation is given by

$$DJ_{\text{vol}}(\Omega)[V] = Dj_1(\Omega)[V] + Dj_2(\Omega)[V] + Dj_3(\Omega)[V]$$

12

(a) Initial configuration with symmetries   (b) Optimal shape

Figure 3: Approximation of a Kelvin cell in a symmetric domain

and the surface formulation by

$$DJ_{\text{surf}}(\Omega)[V] = Dj_4(\Omega)[V].$$

The next section is devoted to the numerical realization of algorithm 2 in order to solve the PDE constrained model problems given above.

4   Numerical results

We focus on three numerical experiments, which are selected in order to demonstrate challenges arising for large-scale multigrid shape optimizations. All results involved are computed at the High Performance Computing Center Stuttgart using the machine HAZELHEN with up to $16\,384$ cores.

In the *first experiment*, we choose $\nu_1 = \nu_2 = 0$ such that we end up with a pure geometric optimization problem without any PDE constraints. Hereby, we want to demonstrate a proper choice of boundary conditions in the mesh deformation in order to obtain results for periodic domains. This optimization problem is motivated by the question for a space-filling cell design with least area of surfaces and only one type of cells. It goes back to the 19th century and is also known as the Kelvin problem, for which he proposed a solution (cf. [25]) based on cells as depicted in figure 3b. His conjecture was disproved by a counter example in [27]. However, these authors propose a solution with two types of cells. The resulting optimal shape might be understood as the result of a biological growing process and used as a building block for finite element models of the human skin (cf. [16]). Figure 3a depicts the initial geometry for the optimization, where we want to exploit symmetries on all outer surfaces. The shape derivative of the objective (2.10) for this particular case is given as the sum of the derivatives (2.53) and (2.54) with $\nu_3 = \nu_4 = 1.0$. Note that this results in a mixture of volume and surface formulations.

The linear and bilinear form defining the shape metric (3.12), which gives also the information for the mesh deformation field U, is chosen to be the weak form of a linear elasticity model. In order to model the periodicity of the domain, we choose the following Dirichlet and Neumann conditions at the outer boundaries depending on the components $U = (U_x, U_y, U_z)^T$:

$$U_x = 0, \quad \sigma\left((0, U_y, U_z)^T\right) \cdot n = 0 \qquad \text{on } \Gamma_{\text{front}} \cup \Gamma_{\text{back}} \qquad (4.1)$$

$$U_y = 0, \quad \sigma\left((U_x, 0, U_z)^T\right) \cdot n = 0 \qquad \text{on } \Gamma_{\text{left}} \cup \Gamma_{\text{right}} \qquad (4.2)$$

$$U_z = 0, \quad \sigma\left((U_x, U_y, 0)^T\right) \cdot n = 0 \qquad \text{on } \Gamma_{\text{bottom}} \cup \Gamma_{\text{top}} \qquad (4.3)$$

<div align="center">(a) Inital shape   (b) Optimal shape</div>

Figure 4: Sharp-edged initial geometry with identical coarse and fine grid under 6 levels of hierarchical refinements and optimal solution on finest level

Figure 5: Cross-section of the tube at 25% of the length for the first iterations with perimeter regularization

This particular choice ensures that nodes are allowed to slide within Γ_{out}, but not to leave the planes. The Lamé parameter are fixed to $\lambda = 0.01$ and $\mu = 0.1$. We apply this shape metric to all numerical examples in this sections, due to the fact that the sliding conditions minimize the influence of the outer boundary on the optimization of Γ_{int}. Especially, when Γ_{int} intersects Γ_{out} like in the volume experiment shown in figure 3, the sliding effect at boundaries is obligatory.

The result of the optimization is visualized in figure 3b. It can be identified as an approximation to a truncated octahedron according to Kelvins conjecture. This shape is also denoted as a tetrakaidecahedron with 6 square faces and 8 hexagonal faces. We obtain this result after 12 gradient steps on a fine grid with approximately $3.9 \cdot 10^6$ and a coarse grid with $60\,533$ finite elements.

Throughout this section all linear systems are solved with the multigrid preconditioned conjugated gradient solver of the software PETSC. Here a symmetric Gauss-Seidel smoother is applied and the coarse grid solver is a direct factorization computed with the SUPERLU-DIST library. This choice is possible, since all PDEs to be solved, i.e., the parabolic diffusion equation and linear elasticity, are discretized as symmetric, positive definite matrices. Moreover, we choose linear finite elements on tetrahedral grids for the discretization of all PDE models involved. Load balancing is performed with the PARMETIS graph partitioning library. As mentioned above, it is sufficient to partition and balance only with respect to tetrahedrons and not also surface triangles at Γ_{int}. This is due to the fact that we plug in the volume form of shape derivatives whenever possible. Thus, there are only very few surface-only operations, which do not dramatically affect the scalability. By solving the deformation equation (3.12) we then obtain a deformation field U. This vector field is added as a deformation to all nodes in the finite element meshes on all multigrid levels. This means that the fine grid solution U is interpolated at the nodes of the coarser grids. The multigrid hierarchy is designed such that for the initial geometry the nodes of the coarse grid are a subset of the nodes of the fine

<div align="center">14</div>

(a) Objective value   (b) Gradient in g^S norm

Figure 6: Objective value and gradient in g^S norm for the pure parabolic test case with regularization in the first 10 iterations

grid. Thus, this property is maintained throughout the whole optimization process. In principle, this means that the shape optimization takes place at the fine grid and the coarse grids are carried along during this process for the preconditioner of the linear solver.

The *second numerical test* is chosen in order to touch upon the main topic of this paper, namely the interaction of multigrid solvers and shape optimization. We are in a simplified situation of the optimization problem (2.11), which is only subject to the diffusion model (2.15)-(2.18). Thus, the objective function is of tracking type with regularization corresponding to $\nu_1 = \nu_3 = 0$. The domain $\Omega = \Omega_{\text{out}} \cup \Omega_{\text{int}}$ has only one inclusion with a jumping diffusion coefficient $k_{\text{out}} = 1.0$ and $k_{\text{int}} = 0.001$. Figure 4 shows this situation. The initial geometry is depicted on the left hand side and the target is shown on the right hand side. Data measurements $\bar{y}$ are generated in terms of the same model equation and a similar shape to the one in figure 4b. We assume two points in time for measurements – one at the final time $T = 15$ and one after 7.5 seconds.

Measurements are assumed to be represented within 10 000 Gaussian type radial basis functions. This involves expensive but necessary computations, since $\bar{y}$ has to be evaluated on varying meshes during the optimization. A general mesh to mesh interpolation on a distributed memory computer can not be computed efficiently, which makes a detour via radial basis functions to the expensive yet scalable method of choice.

The optimization problem is discretized using 7 levels of hierarchical mesh refinements starting from a coarse grid with 3 390 tetrahedral elements to a fine grid with approximately $8.89 \cdot 10^8$. For the diffusion we choose a backward Euler time stepping with a step-size of $\Delta t = 1.5$. In figure 4a, we observe sharp edges in the geometry, which, in principle, should resolve a tube with radius 0.5 and length 5. This is due to the hierarchical grid structure, which is typical for multigrid methods, since coarse and fine grid have to resolve the same geometry. In our particular numerical test, this property of the mesh hierarchy together with the shape metric lead to problems in the discretization.

Figure 5 shows a cross-section of Ω_{int} at 25% of the tubes length during the first iterates of the optimization. We observe that edges, which have a smaller influence on the objective function than planes, tend to remain at their initial position. This effect is intensified by the choice of the shape metric (3.12). Hereby, the impact of the shape derivative (2.52) on the geometry can be understood as a traction on the boundary, which is stronger on large planes compared to sharp edges. The application of a limited memory BFGS update techniques also influences this

15

Figure 7: log-log-plot of L^∞-norm of mean curvature with respect to number of tetrahedral finite elements

behaviour, since relatively large steps are chosen especially in the first iterations.

In figure 5, it can be observed that plain optimization problem tends to overlapping elements. This would lead to an immediate break down of the PDE solver. In order to prevent this, the problem is regularized by the perimeter term j_4 with $\nu_4 = 0.01$ in (2.10). The effect of the perimeter regularization can be observed throughout the figures. As soon as sharp edges are smoothed out by the influence of the perimeter regularization, we can switch off this regularization by setting $\nu_4 = 0$. In our particular case, we do this after 10 quasi-Newton steps. The effect of this strategy is visualized in figure 6. The objective value is depicted on the left hand side, where there is a small step when the regularization is switched off. Note that this step is a bit delayed, since we choose a limited memory BFGS strategy with a storage of 5 gradients. The figure on the right hand side shows the g^S-norm of the gradient, where we observe a jump after 10 iterations. The two figures in 6 can be seen as an indicator for mesh-independent convergence as expected due to the quasi-Newton method. We can thus conclude, that for this particular test, a regularization is only required due to the multigrid approach. Otherwise, if there are no fine-resolved kinks in the initial geometry, one could be successful without regularization.

We should mention that there is a formulation according to [22, proposition 2.50], which is equivalent to (2.54), for the derivative of the perimeter regularization given by

$$Dj_4(\Omega)[V] = \nu_4 \int_\Omega \operatorname{div}(V) - \left\langle \frac{\partial V}{\partial n}, n \right\rangle ds. \tag{4.4}$$

This is attractive from a computational point of view, since the evaluation of κ in each iteration is a surface-only operation. Thus, its scalability is affected by the load balancing, which one would have to perform with respect to both volume and surface elements. However, in our experiments, it seems that for the discretized problem the formulation (2.54) is more successive in eliminating kinks arising due to the multigrid mesh hierarchy. Figure 7 shows the approximation to the mean curvature on the 6 multigrid levels for the initial geometry. The curvature computation follows the approach presented in [11]. An exponential growth of the curvature over the grid levels can be observed, which makes it necessary to adapt ν_4 depending on the finest grid level throughout the refined computations presented in figure 6.

We now concentrate on the *third numerical example*, which is the optimization with respect to the complete system (2.12)-(2.18) gathering diffusion and elasticity. The objective, in our particular case, is then formed by the constants $\nu_1 = 0.15$, $\nu_2 = 0.1$, $\nu_3 = 16.0$ and $\nu_4 = 0.01$. Furthermore, data measurements $\bar{y}$ are chosen to be the simulated values y with the initial geometry. These are again represented in radial basis functions in order to interpolate them on arbitrary meshes. The

16

(a) Initial grid

(b) 10 gradient steps

(c) 20 steps

(d) 30 steps

Figure 8: Optimization with respect to diffusion data measurements and elastic compliance of a cellular structure

first challenge encountered here is that the coarse grid has to resolve the desired geometry as depicted in figure 8. This limits the coarseness, since each inclusion, that should be simulated, has to be present in the entire grid hierarchy. In addition to the problem of large kinks in the fine grid, we now have to deal with coarse grids having a large number of finite elements. Here we want to simulate $7 \cdot 7 \cdot 7$ inclusions resulting in a coarse grid consisting of $74\,096$ tetrahedral finite elements. A coarser grid is not reasonable, since this would have a negative influence on the aspect ratios of the elements, which affects the convergence of iterative solvers. Moreover, the large number of elements in the coarse grid is a major challenge for solvers based on multigrid. In order to be efficient, a direct factorization is usually applied on the coarsest level, which is known to be only scalable to some extent.

Figure 8 shows some snapshots of the optimization on a fine grid with $4.74 \cdot 10^6$ tetrahedral elements. The aim of this particular model is to show how cellular structures can be fortified by tightly stacking cells together close to the surface, where forces are acting. This, to some extent, reflects growing processes in biological skin structures. In this particular case, we are not able to apply BFGS update techniques like in the pure parabolic test case. The quasi-Newton method leads to step sizes which are too large to be feasible as deformations on the finite element grid. We thus apply only a gradient descent method. The geometry after 30 gradient steps, which is depicted in figure 8d, is not yet optimal in the sense that it is a stationary point. Due to increasing aspect ratios of discretization elements, the multigrid

solver does not converge anymore. This illustrates the challenge to choose proper regularization in the case of shape optimization, such that an optimal solution can still be represented in a finite element mesh.

5 Conclusions

One of the main focuses of this paper is to describe the interaction of algorithmic building blocks ranging from a multigrid finite element solver to quasi-Newton methods for shape optimization problems. Here we concentrate on shape spaces and metrics, which are especially suited for large-scale computations. This results in scalable algorithms for supercomputers. Based on complex examples this paper presents the possibilities and challenges of multigrid methods and shape optimization. In particular, it is shown how to handle the increasing values of approximated curvature in an hierarchical grid structure affecting the regularization. The purpose of the underlying models within this article is to form building blocks for further studies of biological cell structures.

Acknowledgment

References

[1] P.A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.

[2] M. Berggren. A unified discrete-continuous sensitivity analysis method for shape optimization. In W. Fitzgibbon et al., editors, *Applied and Numerical Partial Differential Equations*, volume 15 of *Computational Methods in Applied Sciences*, pages 25–39. Springer, 2010.

[3] J. Céa. Conception optimale ou identification de formes calcul rapide de la dérivée directionelle de la fonction coût. *RAIRO Modelisation mathématique et analyse numérique*, 20(3):371–402, 1986.

[4] R. Correa and A. Seeger. Directional derivative of a minmax function. *Nonlinear Analysis: Theory, Methods & Applications*, 9(1):13–22, 1985.

[5] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, volume 22 of *Advances in Design and Control*. SIAM, 2nd edition, 2001.

[6] P. Gangl, A. Laurain, H. Meftahi, and K. Sturm. Shape optimization of an electric motor subject to nonlinear magnetostatics. *SIAM Journal on Scientific Computing*, 37(6):B1002–B1025, 2015.

[7] S. Gross and A. Reusken. *Numerical Methods for Two-Phase Incompressible Flows*, volume 40 of *Computational Mathematics*. Springer, 2010.

[8] J. Haslinger and R.A.E. Mäkinen. *Introduction to Shape Optimization: Theory, Approximation, and Computation*, volume 7 of *Advances in Design and Control*. SIAM, 2003.

[9] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*, volume 15 of *Advances in Design and Control*. SIAM, 2008.

[10] K. Ito, K. Kunisch, and G.H. Peichl. Variational approach to shape derivatives. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(3):517–539, 2008.

[11] M. Meyer, M. Desbrun, P. Schröder, and A.H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In H.-C. Hege and K. Polthier, editors, *Visualization and Mathematics III*, pages 35–57. Springer, 2003.

[12] P.M. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Mathematica*, 10:217–245, 2005.

[13] P.M. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8(1):1–48, 2006.

[14] P.M. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.

[15] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids*. Oxford University Press, 2001.

[16] I. Muha, A. Naegel, S. Stichel, A. Grillo, M. Heisig, and G. Wittum. Effective diffusivity in membranes with tetrakaidekahedral cells and implications for the permeability of human stratum corneum. *Journal of Membrane Science*, 368(1):18–25, 2011.

[17] A. Novotny and J. Sokolowski. *Topological Derivatives in Shape Optimization*. Springer, 2013.

[18] S. Schmidt, E. Wadbro, and M. Berggren. Large-scale three-dimensional acoustic horn optimization. *SIAM Journal on Scientific Computing*, 38(6):B917B940, 2016.

[19] V.H. Schulz, M. Siebenborn, and K. Welker. Efficient PDE constrained shape optimization based on Steklov-Poincaré type metrics. *SIAM Journal on Optimization (accepted)*, 2015.

[20] V.H. Schulz, M. Siebenborn, and K. Welker. Structured inverse modeling in parabolic diffusion problems. *SIAM Journal on Control and Optimization*, 53(6):3319–3338, 2015.

[21] M.J. Sewell. *Maximum and Minimum Principles: A Unified Approach with Applications*. Cambridge University Press, 1987.

[22] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization*, volume 16 of *Computational Mathematics*. Springer, 1992.

[23] K. Sturm. Lagrange method in shape optimization for non-linear partial differential equations: A material derivative free approach. Technical Report No. 1817, Weierstraß-Institut für angewandte Analysis und Stochastik, Berlin, 2013.

[24] K. Sturm. Shape differentiability under non-linear PDE constraints. In A. Pratelli and G. Leugering, editors, *New Trends in Shape Optimization*, volume 166 of *International Series of Numerical Mathematics*, pages 271–300. Springer, 2015.

[25] W. Thomson. On the division of space with minimum partitional area. *Acta Mathematica*, 11(1-4):121–134, 1887.

[26] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, volume 112 of *Applied Mathematics*. American Mathematical Society, 2010.

[27] D. Weaire and R. Phelan. A counter-example to kelvin's conjecture on minimal surfaces. *Philosophical Magazine Letters*, 69(2):107–110, 1994.

[28] K. Welker. *Efficient PDE Constrained Shape Optimization in Shape Spaces*. PhD thesis, Universität Trier, 2016.