Comput Mech (2008) 41:661–681
DOI 10.1007/s00466-007-0223-5

ORIGINAL PAPER

On variational sensitivity analysis and configurational mechanics

Daniel Materna · Franz-Joseph Barthold

Received: 24 March 2007 / Accepted: 12 October 2007 / Published online: 10 November 2007

Abstract This contribution is concerned with the application of variational design sensitivity analysis in the context of structural optimization and configurational mechanics. In both disciplines we consider variations of the material configuration and we use techniques from variational sensitivity analysis in order to solve these problems. We derive the physical and material residual problem in one step by using standard optimization procedures. Furthermore, we investigate the sensitivity of the physical as well as the material residual problem and obtain the coupled saddle point problem based on these sensitivities. Both problems are coupled by the pseudo load operator, which plays an important role by the solution of structural optimization problems. By means of computational examples from mesh optimization and shape optimization, we demonstrate the capability of the proposed theoretical framework.

Keywords Variational sensitivity analysis · Configurational mechanics · Mesh optimization · Shape optimization

1 Introduction

Variational design sensitivity analysis is a branch of structural optimization. The topological connection of bodies, i.e. developing and vanishing holes in structures, are examined in the field of topology optimization whereas the geometrical shape of the construction is considered within shape optimization, see e.g. [17,18,29] and the reference therein for an overview. In these disciplines we consider variations of the material configuration and we are interested in the change of the state variables and the objective functional due to these variations. These sensitivities are required in order to solve the corresponding Lagrangian equation using nonlinear programming algorithms.

In the same manner in configurational mechanics we are interested in changes of the material body and the so-called configurational or material forces are used in the context of material inhomogeneities like crack propagation or phase transitions problems as well as any kind of material defects, see e.g. [26,28,30,38,47,48] and the references therein. The notion of configurational forces was introduced by the fundamental work of Eshelby [21,22]. But apart from the above-mentioned physical inhomogeneities other applications with configurational forces are treated. In a series of recently published papers procedures for mesh optimization based on the use of configurational forces were proposed, see [1, 27,31,42,41,50,51]. The first discussion in the direction of configurational forces caused by discretization was done by [13]. Furthermore, applications for the optimization of truss structures were presented in [2,14].

First steps for the optimization of finite element meshes based on a discrete formulation of energy minimization with respect to changes of the nodal coordinates as well as the deformation were outlined for instance in [15,16,40,49]. In fact, the authors obtained the same discrete indicators for mesh optimization like the above mentioned approach from configurational mechanics, but they have not called them material or configurational forces.

The relations of configurational mechanics and its application to mesh optimization can be obtained by using techniques

D. Materna (✉) · F.-J. Barthold
Chair of Numerical Methods and Information Processing,
University of Dortmund, August-Schmidt-Straße 8,
44227 Dortmund, Germany
e-mail: daniel.materna@uni-dortmund.de

F.-J. Barthold
e-mail: franz-joseph.barthold@uni-dortmund.de

from variational design sensitivity analysis applied to the energy functional of the problem, see [4,5,33,34]. The energy depends on the state variables and the design variables. The design variables specify the shape and/or the topology of the material body. The first sensitivity of the energy functional with respect to changes in the design leads to the well-known weak form of the material or configurational force equilibrium. In the context of sensitivity analysis, we can interpret the configurational forces as the sensitivity of the energy with respect to variations in the design. The second sensitivity of the energy functional provides information about the sensitivity of the energy release rate and yields a tangent operator in order to solve the material motion problem. Furthermore, we investigate the pseudo load operator, which can be derived from the sensitivity of the physical motion problem or from the sensitivity of the material motion problem. The sensitivity of the physical as well as the sensitivity of the material motion problem depend on the structure of the pseudo load operator. Both problems are coupled by this operator and some important properties are studied in this context.

An application of the proposed treatment is the above mentioned optimization of a given finite element mesh. For mesh optimization, we choose simply the nodal coordinates as design variables. The optimal nodal positions (the design) of a given mesh (the reference configuration) can be calculated using standard optimization algorithms. In this context, the configurational forces on the nodes are indicators in which direction the nodes have to move in order to change the energy.

Furthermore, the minimization of the energy is directly related to the minimization of the compliance of the system or equivalently to the maximization of the stiffness. The minimization of the compliance is an interesting and often used objective functional in shape and topology optimization, see e.g. [10,19]. In this context, the configurational forces on the design boundary are indicators in which direction the boundary has to move in order to minimize the compliance or to maximize the structural stiffness.

The paper is organized as follows. After some preliminaries about the notation, we begin with an abstract setting of a classical structural optimization problem and formulate an energy minimization problem. We obtain the classic physical residual problem and the material residual problem in one step. We perform variational sensitivity analysis for the physical as well as for the material residual problem and we obtain variational equations for the sensitivities of the state and the design variables, respectively. With these sensitivities, we can formulate a solution algorithm for the coupled equations. Finally, we summarize the theoretical as well as computational treatment by means of a simple model problem and selected examples from mesh optimization as well as shape optimization.

2 Preliminaries

2.1 Kinematics

We consider an open bounded material body with an undeformed reference configuration $\Omega_R \subset \mathbb{E}^3$ with a piecewise smooth, polyhedral and Lipschitz-continuous boundary $\Gamma = \partial\Omega_R$ such that $\Gamma = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$, where Γ_D denotes the Dirichlet boundary and Γ_N the Neumann boundary, respectively. The corresponding deformed current configuration is denoted by $\Omega_t \subset \mathbb{E}^3$. The deformation of the material body from Ω_R into a deformed configuration Ω_t is given by the nonlinear mapping

$$\boldsymbol{\varphi} : \begin{cases} \Omega_R \times I_t \to \Omega_t \subset \mathbb{E}^3 \\ (\boldsymbol{X}, t) \quad \mapsto \boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t) \end{cases} . \tag{1}$$

Here, $\boldsymbol{\varphi}$ maps the material particle $\boldsymbol{X}$ from the reference configuration Ω_R to the spatial coordinates $\boldsymbol{x}$ in the deformed configuration Ω_t for any fixed time $t \in I_t$. The corresponding deformation gradient, i.e. the tangent map of $\boldsymbol{\varphi}$ from the material tangent space $T_X\Omega_R$ to the spatial tangent space $T_x\Omega_t$, as well as its Jacobian J are given by

$$\boldsymbol{F} = \nabla_X \boldsymbol{\varphi} : T_X\Omega_R \to T_x\Omega_t \quad \text{and} \quad J = \det \boldsymbol{F}. \tag{2}$$

Furthermore, we assume that the deformation is injective, sufficiently smooth and that $J > 0$, such that there exist the inverse deformation mapping

$$\boldsymbol{\phi} : \begin{cases} \Omega_t \times I_t \to \Omega_R \subset \mathbb{E}^3 \\ (\boldsymbol{x}, t) \quad \mapsto \boldsymbol{X} = \boldsymbol{\phi}(\boldsymbol{x}, t) \end{cases} . \tag{3}$$

The corresponding deformation gradient $\boldsymbol{f}$ and its Jacobian j are given by

$$\boldsymbol{f} = \nabla_x \boldsymbol{\phi} : T_x\Omega_t \to T_X\Omega_R \quad \text{and} \quad j = \det \boldsymbol{f}. \tag{4}$$

2.2 Enhanced kinematics

In the most mechanical problems we consider a material body with a fixed reference configuration Ω_R. For problems with changing reference configurations, e.g. shape or topology optimization or problems from configurational mechanics, it is useful to extend the standard kinematics.

The above introduced kinematical settings in the referential placement Ω_R and the current placement Ω_t could be enhanced by using the *intrinsic formulation* by Noll [46]. This is based on the concept of a *differentiable manifold*, see for instance [12,32,52] for details.

An enhancement of the above mentioned approach by Noll in the context of variational design sensitivity analysis was proposed by [3,6,8]. Following the intrinsic concept, a given manifold can be described locally using an intrinsic coordinate system defined on an independent continuous parameter space P_Θ with local coordinates $\boldsymbol{\Theta}$. Without going

into detail, this leads to two fundamental mappings, a design dependent *local reference placement mapping*

$$\boldsymbol{\kappa} : \begin{cases} P_\Theta \times I_s \to \Omega_R \subset \mathbb{E}^3 \\ (\boldsymbol{\Theta}, s) \quad \mapsto \boldsymbol{X} = \boldsymbol{\kappa}(\boldsymbol{\Theta}, s) \end{cases} \tag{5}$$

and a time dependent *local current placement mapping*

$$\boldsymbol{\mu} : \begin{cases} P_\Theta \times I_t \to \Omega_t \subset \mathbb{E}^3 \\ (\boldsymbol{\Theta}, t) \quad \mapsto \boldsymbol{x} = \boldsymbol{\mu}(\boldsymbol{\Theta}, t) \end{cases} \tag{6}$$

for any fixed time $t \in I_t$ and any design $s \in I_s$.[1] The corresponding tangent maps and its Jacobians are given by

$$\boldsymbol{K} = \nabla_\Theta \boldsymbol{\kappa} : T_\Theta P_\Theta \to T_X \Omega_R \quad J_K = \det \boldsymbol{K} \tag{7}$$

$$\boldsymbol{M} = \nabla_\Theta \boldsymbol{\mu} : T_\Theta P_\Theta \to T_x \Omega_t \quad J_M = \det \boldsymbol{M}. \tag{8}$$

With these mappings, the deformation map $\boldsymbol{\varphi}$ (1) and its tangent map can be written in the form

$$\boldsymbol{\varphi} = \boldsymbol{\mu} \circ \boldsymbol{\kappa}^{-1} \quad \text{and} \quad \boldsymbol{F} = \nabla_X \boldsymbol{\varphi} = \boldsymbol{M}\boldsymbol{K}^{-1} \tag{9}$$

and for the inverse mapping $\boldsymbol{\phi}$ (3) follows

$$\boldsymbol{\phi} = \boldsymbol{\kappa} \circ \boldsymbol{\mu}^{-1} \quad \text{and} \quad \boldsymbol{f} = \nabla_x \boldsymbol{\phi} = \boldsymbol{K}\boldsymbol{M}^{-1}, \tag{10}$$

respectively. The difference vector between the reference and current placements is the displacement $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{X}$. The corresponding *local displacement mapping* $\boldsymbol{\upsilon}$ written in terms of the local mappings $\boldsymbol{\kappa}$ and $\boldsymbol{\mu}$ is given with

$$\boldsymbol{u} = \boldsymbol{\upsilon}(\boldsymbol{\Theta}, s, t) = \boldsymbol{\mu}(\boldsymbol{\Theta}, t) - \boldsymbol{\kappa}(\boldsymbol{\Theta}, s). \tag{11}$$

Remark 1 The *local reference placement mapping* is parameterized by a design parameter s, i.e. $\boldsymbol{X} = \boldsymbol{\kappa}(\boldsymbol{\Theta}, s)$. With this in mind, we introduce in an abstract sense a *generalized design* or *control function* $\boldsymbol{s} \in \mathcal{D}$, which specifies the current reference configuration Ω_R, i.e. $\Omega_R = \Omega_R(\boldsymbol{s})$ or

$$\boldsymbol{X} = f(\boldsymbol{s}). \tag{12}$$

Here, $\mathcal{D}$ denotes the space with all admissible design or control functions. Therefore, the material particle $\boldsymbol{X}$ as well as the spatial particle $\boldsymbol{x}$ depend on the design variable, i.e. $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{s})$ and $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X}(\boldsymbol{s}))$. The explicit coupling between $\boldsymbol{X}$ and $\boldsymbol{s}$, i.e. the explicit structure of $f(\boldsymbol{s})$ depends on the particular problem. The *local current placement mapping* $\boldsymbol{x} = \boldsymbol{\mu}(\boldsymbol{\Theta}, t)$ can be expressed by the *local displacement mapping* $\boldsymbol{u} = \boldsymbol{\upsilon}(\boldsymbol{\Theta}, s, t)$. With these, a quantity $(\cdot)$, which depends on the state variable $\boldsymbol{u}$ and the design variable $\boldsymbol{s}$ is denoted by $(\cdot)(\boldsymbol{u}, \boldsymbol{s})$.

[1] Here, s is used as a general scalar (time-like) design variable, which parameterizes in an abstract sense the material body in the reference configuration Ω_R, i.e. $\Omega_R = \Omega_R(s)$. This could be a parametrization for the material points $\boldsymbol{X} = \boldsymbol{X}(s)$, density $\rho = \rho(s)$, mass $m = m(s)$, material properties etc. In this paper, we consider only an abstract parametrization for the material points, see Remark 1.

For problems with changes in the material configuration, e.g. shape optimization or problems from configurational mechanics, the intrinsic formulation in local coordinates has advantages, because we deal with two independent placement mappings, see [3,6,8] for more details. The relationship of this formalism to the arbitrary Lagrangian-Eulerian (ALE) approach is discussed in [6].

2.3 Variations and derivatives

The variation of a quantity $(\cdot)(\boldsymbol{\varphi}; \boldsymbol{X}, t)$ with respect to $\boldsymbol{x}$ at fixed $\boldsymbol{X}$ and t is given by

$$\delta_x(\cdot) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(\cdot)(\boldsymbol{\varphi} + \varepsilon\delta\boldsymbol{\varphi}; \boldsymbol{X}, t)\bigg|_{\varepsilon=0} =: (\cdot)'_x \tag{13}$$

and the variation of $(\cdot)(\boldsymbol{\phi}; \boldsymbol{x}, t)$ with respect to $\boldsymbol{X}$ at fixed $\boldsymbol{x}$ and t is

$$\delta_X(\cdot) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(\cdot)(\boldsymbol{\phi} + \varepsilon\delta\boldsymbol{\phi}; \boldsymbol{x}, t)\bigg|_{\varepsilon=0} =: (\cdot)'_X. \tag{14}$$

Remark 2 In order to avoid confusion between the small $\boldsymbol{x}$ and the capital $\boldsymbol{X}$ and due to the fact that the reference configuration of the material body is prescribed by a design function $\boldsymbol{s}$, we prefer to use $\delta_u(\cdot)$ instead of $\delta_x(\cdot)$ and $\delta_s(\cdot)$ instead of $\delta_X(\cdot)$ respectively. Furthermore, the short notation $(\cdot)'$ denote the total variation of a quantity $(\cdot)$ and $(\cdot)'_u$ as well as $(\cdot)'_s$ the partial variation with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$, i.e.

$$(\cdot)'_u := (\cdot)'_x = \delta_x(\cdot) \quad \text{and} \quad (\cdot)'_s := (\cdot)'_X = \delta_X(\cdot). \tag{15}$$

With this notation, the total variation of a quantity $(\cdot)(\boldsymbol{u}, \boldsymbol{s})$, which depends on the deformation and the design, is given by the partial variation with respect to $\boldsymbol{u}$ at a fixed design $\hat{\boldsymbol{s}}$ as well as the partial variation with respect to $\boldsymbol{s}$ at a fixed deformation $\hat{\boldsymbol{u}}$, i.e.

$$(\cdot)'(\boldsymbol{u}, \boldsymbol{s}) = (\cdot)'_u(\boldsymbol{u}, \hat{\boldsymbol{s}}) + (\cdot)'_s(\hat{\boldsymbol{u}}, \boldsymbol{s}). \tag{16}$$

In the same way we define the second variations $(\cdot)''_{uu} := \delta^2_{xx}(\cdot)$ and $(\cdot)''_{ss} := \delta^2_{XX}(\cdot)$ as well as the mixed variations $(\cdot)''_{us} := \delta^2_{xX}(\cdot)$ and $(\cdot)''_{su} := \delta^2_{Xx}(\cdot)$, respectively.

Furthermore, we introduce a distinction between the total partial derivative $D_s(\cdot)$ and the explicit partial derivative $\partial_s(\cdot)$ of a quantity $(\cdot)$ with respect to a variable $\boldsymbol{s}$. The total partial derivative of a function $f(\boldsymbol{s}, \boldsymbol{u}) = f(\boldsymbol{s}, \boldsymbol{u}(\boldsymbol{s}))$ with respect to $\boldsymbol{s}$ and the explicit partial derivative are connected by the relation

$$D_s f(\boldsymbol{s}, \boldsymbol{u}) = \partial_s f + \frac{\partial f}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{s}}. \tag{17}$$

3 A structural optimization problem

3.1 Abstract setting

In structural optimization we are interested in the solution of the following problem:

Problem 1 Find $\{\boldsymbol{u}, \boldsymbol{s}\} \in \mathcal{V} \times \mathcal{D}$ of the objective functional $\mathcal{J} : \mathcal{V} \times \mathcal{D} \to \mathbb{R}$ such that

$$\mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) \to \min_{\boldsymbol{u},\boldsymbol{s} \in \mathcal{V}\times\mathcal{D}} \tag{18}$$

subject to the constraint

$$\boldsymbol{A}(\boldsymbol{u}, \boldsymbol{s}) = \boldsymbol{0} \quad \text{in } \Omega, \tag{19}$$

where $\boldsymbol{A}$ is an elliptic differential operator for the state function $\boldsymbol{u} \in \mathcal{V}$ and $\boldsymbol{s} \in \mathcal{D}$ is a generalized design function. Additionally we have to fulfil boundary conditions for the state and for the design, i.e.

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \Gamma_{D_u} \quad \text{and} \quad \boldsymbol{s} = \bar{\boldsymbol{s}} \quad \text{on } \Gamma_{D_s}. \tag{20}$$

Here, $\mathcal{V}$ denotes the usual Sobolev space and $\mathcal{D}$ the space with all admissible designs. Furthermore, Γ_{D_u} denotes the Dirichlet boundary for the state variables and the Γ_{D_s} the corresponding boundary for the design variables.

To solve this problem we introduce the corresponding Lagrangian functional

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\lambda}) :=\ & \mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) + \langle \boldsymbol{A}(\boldsymbol{u}, \boldsymbol{s}), \boldsymbol{\lambda} \rangle \\ =\ & \mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) + \mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\lambda}),\end{aligned} \tag{21}$$

where

$$\mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\lambda}) := a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\lambda}) - F(\boldsymbol{\lambda}) = 0 \quad \forall\, \boldsymbol{\lambda} \in \mathcal{V} \tag{22}$$

is the weak form of the differential operator $\boldsymbol{A}$ written in terms of the physical residual $\mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \cdot)$. We seek for stationary points of $\mathcal{L}$ which are candidates for optimal solutions of the system

$$\mathcal{L}'(\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\lambda})(\boldsymbol{\eta}, \boldsymbol{\psi}, \boldsymbol{\nu}) := \nabla \mathcal{L} = \frac{\partial \mathcal{L}(\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\lambda})}{\partial (\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\lambda})} = 0.$$

This is a boundary value problem for the triple $\{\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\lambda}\} \in \mathcal{V} \times \mathcal{D} \times \mathcal{V}$,

$$\begin{aligned}\mathcal{L}' = \left\{ \begin{array}{l} \mathcal{J}'_u(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) + \mathcal{R}'_u(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\lambda}, \boldsymbol{\eta}) \\ \mathcal{J}'_s(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) + \mathcal{R}'_s(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\lambda}, \boldsymbol{\psi}) \\ \mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\nu}) \end{array} \right\} = \boldsymbol{0} \\ \forall\, \{\boldsymbol{\eta}, \boldsymbol{\psi}, \boldsymbol{\nu}\} \in \mathcal{V} \times \mathcal{D} \times \mathcal{V},\end{aligned} \tag{23}$$

where $\mathcal{J}'_u, \mathcal{J}'_s, \mathcal{R}'_u, \mathcal{R}'_s$ are the partial variations of the objective functional $\mathcal{J}$ and the physical residual $\mathcal{R}$ with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$. The last equation of (23) is simply the physical residual (22) for the state $\boldsymbol{u}$.

Remark 3 For the solution $\boldsymbol{u} \in \mathcal{V}$ of (22) with a fixed $\hat{\boldsymbol{s}}$, the corresponding adjoint or dual variable $\boldsymbol{\lambda} \in \mathcal{V}$ is the solution of the so-called adjoint or dual problem Eq. $(23)_1$ at the current linearization point $\hat{\boldsymbol{u}}$, i.e. $\boldsymbol{\lambda}$ is the solution of

$$k(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\lambda}, \boldsymbol{\eta}) = -\mathcal{J}'_u(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\eta}) \quad \forall\, \boldsymbol{\eta} \in \mathcal{V}. \tag{24}$$

The bilinear form $k(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\lambda}, \boldsymbol{\eta}) := \mathcal{R}'_u(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\lambda}, \boldsymbol{\eta})$ is the tangent stiffness operator at the current linearization point for the solution of the primal problem.

3.2 A special objective function

In general, the objective function is arbitrary and depends on the particular application. As a special case we choose the potential energy of a hyperelastic body as objective function because it has a clear physical meaning and is an interesting objective in mechanics and engineering. The energy is directly related to the compliance or stiffness of the structure and is used as objective functional in shape and topology optimization, see e.g. [10,19]. Let

$$\begin{aligned}E(\boldsymbol{u}, \boldsymbol{s}) :=\ & \int_{\Omega_R} U_R(\boldsymbol{X}; \boldsymbol{\varphi}, \boldsymbol{F})\, \mathrm{d}\Omega \\ =\ & \int_{\Omega_t} U_t(\boldsymbol{x}; \boldsymbol{\phi}, \boldsymbol{f})\, \mathrm{d}\Omega \\ =\ & \int_{P_\Theta} U_\Theta(\boldsymbol{\Theta}; \boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{M}, \boldsymbol{K})\, \mathrm{d}\Omega\end{aligned} \tag{25}$$

be the total potential energy, where U_R is the energy density in Ω_R, U_t the energy written in terms of the inverse deformation in Ω_t and U_Θ is the energy density defined on a continuous parameter space P_Θ. Here, $U = W + V$ contain the stored strain energy W and external potential energy V are each defined on the different configurations. We assume that the body forces are conservative forces.

With these assumptions, the Problem 1 can be reformulated in the following form:

Problem 2 Find $\{\boldsymbol{u}, \boldsymbol{s}\} \in \mathcal{V} \times \mathcal{D}$ of the objective function $\mathcal{J} : \mathcal{V} \times \mathcal{D} \to \mathbb{R}$ such that

$$\mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) = E(\boldsymbol{u}, \boldsymbol{s}) \to \min_{\{\boldsymbol{u},\boldsymbol{s}\} \in \mathcal{V}\times\mathcal{D}} \tag{26}$$

subject to the constraints

$$\boldsymbol{A}(\boldsymbol{u}, \boldsymbol{s}) = \boldsymbol{0} \quad \text{in } \Omega, \quad \boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \Gamma_{D_u}, \quad \boldsymbol{s} = \bar{\boldsymbol{s}} \quad \text{on } \Gamma_{D_s}.$$

For this problem, we have $\mathcal{J}'_u = E'_u = \mathcal{R} = 0$ and the adjoint equation (24) can be rewritten in the form

$$k(\boldsymbol{u}, \hat{\boldsymbol{s}}; \boldsymbol{\lambda}, \boldsymbol{\eta}) = -\mathcal{J}'_u(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\eta}) = -\mathcal{R} = 0 \tag{27}$$

and therefore the Lagrangian variable $\boldsymbol{\lambda}$ becomes zero as well as the Lagrangian (21) remains $\mathcal{L}(\boldsymbol{u}, \boldsymbol{s}) = \mathcal{J}(\boldsymbol{u}, \boldsymbol{s})$. Furthermore, the system (23) is reduced to the following variational problem.

Problem 3 Find $\{\boldsymbol{u}, \boldsymbol{s}\} \in \mathcal{V} \times \mathcal{D}$ such that

$$\mathcal{L}'(\boldsymbol{u}, \boldsymbol{s})(\boldsymbol{\eta}, \boldsymbol{\psi}) = \left\{ \begin{array}{c} \mathcal{J}'_u(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) \\ \mathcal{J}'_s(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) \end{array} \right\} = \left\{ \begin{array}{c} \mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) \\ \mathcal{G}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) \end{array} \right\} = \boldsymbol{0} \quad \forall \{\boldsymbol{\eta}, \boldsymbol{\psi}\} \in \mathcal{V} \times \mathcal{D}\,. \tag{28}$$

The partial variation of J with respect to $\boldsymbol{u}$ leads to the physical residual (22) $\mathcal{R} : \mathcal{V} \to \mathbb{R}$

$$\mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) := a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) - F(\boldsymbol{s}; \boldsymbol{\eta}). \tag{29}$$

In the same manner, variation with respect to changes in the design $\boldsymbol{s}$ leads to the material residual $\mathcal{G} : \mathcal{D} \to \mathbb{R}$ in the form

$$\mathcal{G}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) := b(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) - L(\boldsymbol{s}; \boldsymbol{\psi}). \tag{30}$$

Remark 4 The material residual is also referred to as the weak form of the material or configurational forces equilibrium as well as the weak form of the pseudo-momentum equation, [26,30,37,39]. In the case of a homogeneous elastic body the material residual is the weak form of the inverse deformation problem. For a detailed discussion about the direct and inverse deformation problem and its duality see for instance [24,25,27,31,47].

The semilinear forms $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ and $b : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ contain the parts of the partial variations with respect to deformation gradients, i.e.

$$\begin{aligned} a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) &:= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{P_\Theta} U_\Theta(\boldsymbol{\Theta}; \boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{M} + \varepsilon\, \nabla\tilde{\boldsymbol{\eta}}, \boldsymbol{K})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{\Omega_R} U_R(\boldsymbol{X}; \boldsymbol{\varphi}, \boldsymbol{F} + \varepsilon\, \nabla\boldsymbol{\eta})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \end{aligned} \tag{31}$$

$$\begin{aligned} b(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) &:= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{P_\Theta} U_\Theta(\boldsymbol{\Theta}; \boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{M}, \boldsymbol{K} + \varepsilon\, \nabla\tilde{\boldsymbol{\psi}})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{\Omega_t} U_t(\boldsymbol{x}; \boldsymbol{\phi}, \boldsymbol{f} + \varepsilon\, \nabla\boldsymbol{\psi})\, \mathrm{d}\Omega \Big|_{\varepsilon=0}. \end{aligned} \tag{32}$$

The linear functionals $F : \mathcal{V} \to \mathbb{R}$ and $L : \mathcal{D} \to \mathbb{R}$ contain the parts of the partial variations with respect to deformations

$$\begin{aligned} F(\boldsymbol{s}; \boldsymbol{\eta}) &:= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{P_\Theta} U_\Theta(\boldsymbol{\Theta}; \boldsymbol{\mu} + \varepsilon\, \tilde{\boldsymbol{\eta}}, \boldsymbol{\kappa}, \boldsymbol{M}, \boldsymbol{K})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{\Omega_R} U_R(\boldsymbol{X}; \boldsymbol{\varphi} + \varepsilon\, \boldsymbol{\eta}, \boldsymbol{F})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \end{aligned} \tag{33}$$

$$\begin{aligned} L(\boldsymbol{s}; \boldsymbol{\psi}) &:= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{P_\Theta} U_\Theta(\boldsymbol{\Theta}; \boldsymbol{\mu}, \boldsymbol{\kappa} + \varepsilon\, \tilde{\boldsymbol{\psi}}, \boldsymbol{M}, \boldsymbol{K})\, \mathrm{d}\Omega \Big|_{\varepsilon=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_{\Omega_t} U_t(\boldsymbol{x}; \boldsymbol{\phi} + \varepsilon\, \boldsymbol{\psi}, \boldsymbol{f})\, \mathrm{d}\Omega \Big|_{\varepsilon=0}. \end{aligned} \tag{34}$$

Using standard pull back and push forward operations we can transform all quantities into the different domains. For details about the kinematical settings and the variational techniques for shape sensitivity analysis based on the intrinsic formulation in local coordinates see [3,6,8].

The semilinear form $b(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi})$ leads after a pull back to the reference configuration to

$$b(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) = \int_{\Omega_R} (\boldsymbol{\Sigma} + V_R \boldsymbol{1}) : \nabla_X \boldsymbol{\psi}\, \mathrm{d}\Omega \tag{35}$$

where $\boldsymbol{\Sigma}$ is the well-known energy momentum or Eshelby tensor [21,22].

Remark 5 In the context of design sensitivity analysis, we can interpret $\boldsymbol{\Sigma}$ as the configurational forces associated to the variation in the design (the configuration) or rather they are the sensitivities of the energy functional with respect to variations in the design [4,5,33,34]. Problems from configurational mechanics, which are based on hyperelasticity, i.e. which are not deformation path dependent, can be interpreted as shape optimization problems.

Remark 6 In the nonlinear case the operator $a(\boldsymbol{u}, \boldsymbol{s}; \cdot)$ is a semilinear form, i.e. only linear with respect to all arguments on the right of the semicolon. In the linear case $a(\boldsymbol{u}, \boldsymbol{s}; \cdot)$ becomes a bilinear form. By contrast, the material problem is always strongly nonlinear, even if the physical problem is linear. This follows from the fact, that the inverse deformation gradient $\boldsymbol{f} = \nabla\boldsymbol{\phi} = \nabla(\boldsymbol{\varphi}^{-1})$ is nonlinear, even if the direct deformation $\boldsymbol{\varphi}$ is linear.

An explicit formulation of the above expressions for a simple model problem is given in Appendix A.

3.3 Sensitivity of the objective functional

In structural optimization we are interested in the change of the objective functional and the constrains due to variations in the design. In the context of structural optimization this is termed as *design sensitivity analysis*.

With the total partial derivative (17), we introduce the total partial variation of a functional $\mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) = \mathcal{J}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s})$ with respect to $\boldsymbol{s}$ in the form

$$D_s\mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) \cdot \delta\boldsymbol{s} = \partial_s\mathcal{J} \cdot \delta\boldsymbol{s} + \frac{\partial\mathcal{J}}{\partial\boldsymbol{u}} \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{s}} \cdot \delta\boldsymbol{s}. \tag{36}$$

The above sensitivity relation depends on the partial variation $\partial_s\mathcal{J} \cdot \delta\boldsymbol{s} = \mathcal{J}'_s(\boldsymbol{u}, \boldsymbol{s}, \delta\boldsymbol{s})$ and a second part with the derivatives $\partial_u\mathcal{J}$ and the changes in the state $\partial_s\boldsymbol{u}$, which is a result of the functional dependencies. The derivatives $\partial_s\mathcal{J}$ and $\partial_u\mathcal{J}$ are given explicitly, but the derivative of the deformation $\partial_s\boldsymbol{u}$ with respect to $\boldsymbol{s}$ is given implicitly, because they are the results of an analysis for the deformation, see Sect. 3.4. This is

a classical well-known expression in structural optimization for the design sensitivity of the objective functional.

In the case of $\mathcal{J}(\boldsymbol{u},\boldsymbol{s}) = E(\boldsymbol{u},\boldsymbol{s})$, i.e. we choose the energy as objective functional, we have

$$\mathcal{J}'_u = E'_u = \partial_u \mathcal{J} \cdot \delta\boldsymbol{u} = \mathcal{R}(\boldsymbol{u},\boldsymbol{s},\delta\boldsymbol{u}) = 0 \tag{37}$$

and $\mathcal{J}'_s = \partial_s \mathcal{J} \cdot \delta\boldsymbol{s} = \mathcal{G}(\boldsymbol{u},\boldsymbol{s},\delta\boldsymbol{s})$. With this, the second part of the above sensitivity relation vanishes and it remains only the material residual, i.e.

$$D_s\mathcal{J}(\boldsymbol{u}(\boldsymbol{s});\boldsymbol{s}) \cdot \delta\boldsymbol{s} = \mathcal{J}'_s(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{s}) = \mathcal{G}(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{s}). \tag{38}$$

Therefore, in the context of structural optimization, we can interpret the material residual or configurational forces as the sensitivity of the energy with respect to variations in the design, see Remark 5.

3.4 Sensitivity of the physical residual

The solution of the Lagrangian requires the partial variations of the physical residual with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$, respectively. The total variation of the physical residual $\mathcal{R} = a(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta}) - F(\boldsymbol{s};\boldsymbol{\eta}) = 0$ reads

$$\mathcal{R}' = \mathcal{R}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) + \mathcal{R}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) = 0 \tag{39}$$

where the partial variations are given by

$$\mathcal{R}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) := a'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) - F'_u(\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) \tag{40}$$

$$\mathcal{R}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) := a'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) - F'_s(\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}). \tag{41}$$

Note, that $\mathcal{R}'_u = \mathcal{J}''_{uu} = E''_{uu}$ and $\mathcal{R}'_s = \mathcal{J}''_{us} = E''_{us}$, respectively.

We introduce for the variations of the physical residual $\mathcal{R}$ with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$ the operators

$$k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) := \mathcal{R}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) \tag{42}$$

and

$$p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) := \mathcal{R}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) \tag{43}$$

where $k(\boldsymbol{u},\boldsymbol{s};\cdot,\cdot)$ is the well-known *tangent physical stiffness operator* and we call $p(\boldsymbol{u},\boldsymbol{s};\cdot,\cdot)$ the *tangent pseudo or fictitious load operator* for the physical problem. Both operators are bilinear forms $k : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ and $p : \mathcal{V}\times\mathcal{D} \to \mathbb{R}$.

Remark 7 In earlier publications the operator $p(\cdot,\cdot)$ was denoted by $s(\cdot,\cdot)$ and called *tangent sensitivity*, see [3,6,8,33,34]. In order to avoid confusions, it is much more consistent to call $p(\cdot,\cdot)$ the pseudo load operator.

With these notations the total variation yields the form

$$\mathcal{R}' = k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) + p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s}) = 0. \tag{44}$$

After rearranging the above terms we can formulate the following sensitivity equation for the physical problem.

Problem 4 Let $\delta\hat{\boldsymbol{s}} \in \mathcal{D}$ be a given fixed design variation. Find $\delta\boldsymbol{u} \in \mathcal{V}$ such that

$$k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) = -Q(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\hat{\boldsymbol{s}}) \quad \forall\,\boldsymbol{\eta} \in \mathcal{V}, \tag{45}$$

where

$$Q(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\hat{\boldsymbol{s}}) := p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\hat{\boldsymbol{s}}) = \mathcal{R}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\hat{\boldsymbol{s}}) \tag{46}$$

is the *pseudo load* of the physical problem for the variation $\delta\hat{\boldsymbol{s}}$.

This is a variational equation for the sensitivity of the deformation due to changes in the design. For a given variation in the design $\delta\hat{\boldsymbol{s}}$, we can calculate the variation in the state $\delta\boldsymbol{u}$.

Remark 8 In general the pseudo load operator $p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s})$ is a bilinear form $p : \mathcal{V}\times\mathcal{D} \to \mathbb{R}$. For a chosen fixed $\delta\hat{\boldsymbol{s}}$ it becomes a linear functional $Q : \mathcal{V} \to \mathbb{R}$ and is called *pseudo load*, because it plays the role of a *load* in the sensitivity equation (45) and is denoted by $Q(\boldsymbol{u},\boldsymbol{s};\cdot,\delta\hat{\boldsymbol{s}})$, i.e. $Q(\boldsymbol{u},\boldsymbol{s};\cdot,\delta\hat{\boldsymbol{s}}) = p(\boldsymbol{u},\boldsymbol{s};\cdot,\delta\hat{\boldsymbol{s}})$.

3.5 Sensitivity of the material residual

In the same way we can perform the total variation of the material residual $\mathcal{G} = b(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi}) - L(\boldsymbol{s};\boldsymbol{\psi}) = 0$. This reads

$$\mathcal{G}' = \mathcal{G}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) + \mathcal{G}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) = 0 \tag{47}$$

where the partial variations are given by

$$\mathcal{G}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) := b'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) - L'_u(\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) \tag{48}$$

$$\mathcal{G}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) := b'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) - L'_s(\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}). \tag{49}$$

Note, that $\mathcal{G}'_u = \mathcal{J}''_{su} = E''_{su}$ and $\mathcal{G}'_s = \mathcal{J}''_{ss} = E''_{ss}$, respectively.

We introduce for the variations of the material residual $\mathcal{G}$ with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$ the operators

$$d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) := \mathcal{G}'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) \tag{50}$$

and

$$t(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) := \mathcal{G}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) \tag{51}$$

where $d(\cdot;\cdot)$ is the so-called *tangent material stiffness* or *tangent material operator* in order to highlight the duality to the tangent physical stiffness (42) and we call $t(\cdot;\cdot)$ the *tangent pseudo or fictitious load operator* for the material problem, compare with Eq. (43). Both operators are bilinear forms $d : \mathcal{D}\times\mathcal{D} \to \mathbb{R}$ and $t : \mathcal{D}\times\mathcal{V} \to \mathbb{R}$.

As a result of the permutableness of variations, i.e.

$$\mathcal{G}'_u = \mathcal{J}''_{su} = \mathcal{J}''_{us} = \mathcal{R}'_s \tag{52}$$

and the symmetry of the bilinear forms, i.e. $a'_s(\boldsymbol{u},\boldsymbol{s};\cdot,\boldsymbol{\eta}) = a'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\cdot)$, we obtain for the variation of $\mathcal{G}$ with respect to $\boldsymbol{u}$

$$\begin{aligned}\mathcal{G}'_u(\boldsymbol{u},\boldsymbol{s};\cdot,\delta\boldsymbol{u}) &= \mathcal{R}'_s(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\cdot) = p(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\cdot)\\ &= a'_s(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\cdot) - F'_s(\boldsymbol{s};\delta\boldsymbol{u},\cdot).\end{aligned} \tag{53}$$

Thus, due to symmetry, the partial variation $\mathcal{G}'_u$ leads to the tangent pseudo load operator of the physical problem (43), i.e.

$$t(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u}) = p(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\boldsymbol{\psi}). \tag{54}$$

Therefore, we have additional to specify only the material tangent operator $d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) = \mathcal{G}'_s$.

With these notations the total variation yields the form

$$\mathcal{G}' = p(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\boldsymbol{\psi}) + d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) = 0. \tag{55}$$

After rearranging the above terms we can formulate the following sensitivity equation for the material problem.

Problem 5 Let $\delta\hat{\boldsymbol{u}} \in \mathcal{V}$ be a given fixed variation in the state. Find $\delta\boldsymbol{s} \in \mathcal{D}$ such that

$$d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) = -Q(\boldsymbol{u},\boldsymbol{s};\delta\hat{\boldsymbol{u}},\boldsymbol{\psi}) \quad \forall\,\boldsymbol{\psi} \in \mathcal{D}, \tag{56}$$

where

$$Q(\boldsymbol{u},\boldsymbol{s};\delta\hat{\boldsymbol{u}},\boldsymbol{\psi}) := p(\boldsymbol{u},\boldsymbol{s};\delta\hat{\boldsymbol{u}},\boldsymbol{\psi}) = \mathcal{G}'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\hat{\boldsymbol{u}}) \tag{57}$$

is the *pseudo load* of the material problem for the variation $\delta\hat{\boldsymbol{u}}$.

This is a variational equation for the sensitivity of the design due to changes in the deformation. For a given variation in the state $\delta\hat{\boldsymbol{u}}$, we can calculate the variation in the design $\delta\boldsymbol{s}$.

Remark 9 It is interesting to note, that due to symmetry both the sensitivity of the deformation and the sensitivity of the design depend on the pseudo load operator $p(\cdot,\cdot)$. Therefore, this operator plays an important role for the solution of the minimization problem for $\boldsymbol{u}$ and $\boldsymbol{s}$.

Remark 10 For a homogeneous elastic body, the material residual or configurational forces are directly related to the well-known J-integral, which is in linear fracture mechanics equal to the energy release rate G, i.e.

$$J = G = -\mathcal{G}(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi}). \tag{58}$$

Therefore, the sensitivity of the material residual (55) can be used to calculate the sensitivity of the J-integral or energy release rate G, which can be derived as

$$G' = -(d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s}) + p(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\boldsymbol{\psi})). \tag{59}$$

In order to obtain a dependency only from the variation in the design $\delta\boldsymbol{s}$, we can substitute the sensitivity of the deformation (45) to eliminate $\delta\boldsymbol{u}$, see Remark 15 for the discrete approximation. An application for the sensitivity of the energy release rate of a problem from fracture mechanics is given in [7].

3.6 Solution algorithm with a full Newton method

We apply the Newton method on the continuous level in order to get the solution of (28). A Taylor expansion with $\boldsymbol{y} := \{\boldsymbol{u},\boldsymbol{s}\}^{\mathrm{T}}$ reads

$$\mathcal{L}'(\boldsymbol{y}+\Delta\boldsymbol{y}) = \mathcal{L}'(\boldsymbol{y}) + \nabla\mathcal{L}'(\boldsymbol{y})(\Delta\boldsymbol{y}) + \mathcal{O} = \boldsymbol{0}\,. \tag{60}$$

The remainder $\mathcal{O}$ contains higher order terms which can usually be neglected. Each Newton step requires the solution of the linear system

$$\boldsymbol{H}(\boldsymbol{y})(\Delta\boldsymbol{y}) = \begin{Bmatrix} \mathcal{R}'_u & \mathcal{R}'_s \\ \mathcal{G}'_u & \mathcal{G}'_s \end{Bmatrix} = -\mathcal{L}'(\boldsymbol{y}), \tag{61}$$

where $\boldsymbol{H}(\boldsymbol{y})(\Delta\boldsymbol{y}) := \nabla\mathcal{L}'(\boldsymbol{y})(\Delta\boldsymbol{y})$ is the Hessian matrix of the Lagrangian $\mathcal{L}(\boldsymbol{u},\boldsymbol{s})$. The Hessian contains the partial variations of $\mathcal{R}$ and $\mathcal{G}$, i.e.

$$\mathcal{R}'_u = a'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{u}) - F'_u(\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{u}) \tag{62}$$

$$\mathcal{R}'_s = a'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{s}) - F'_s(\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{s}) \tag{63}$$

$$\mathcal{G}'_u = b'_u(\boldsymbol{u},\boldsymbol{s};\Delta\boldsymbol{u},\boldsymbol{\psi}) - L'_u(\boldsymbol{s};\Delta\boldsymbol{u},\boldsymbol{\psi}) \tag{64}$$

$$\mathcal{G}'_s = b'_s(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\Delta\boldsymbol{s}) - L'_s(\boldsymbol{s};\boldsymbol{\psi},\Delta\boldsymbol{s}). \tag{65}$$

With the definitions of the sensitivities of the physical (44) and material residual (55)

$$\begin{aligned}\mathcal{R}' &= \mathcal{R}'_u + \mathcal{R}'_s = k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) + p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{s})\\ \mathcal{G}' &= \mathcal{G}'_u + \mathcal{G}'_s = p(\boldsymbol{u},\boldsymbol{s};\delta\boldsymbol{u},\boldsymbol{\psi}) + d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s})\end{aligned}$$

follows the Hessian

$$\boldsymbol{H}(\boldsymbol{y})(\Delta\boldsymbol{y}) := \begin{Bmatrix} k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{u}) & p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{s}) \\ p(\boldsymbol{u},\boldsymbol{s};\Delta\boldsymbol{u},\boldsymbol{\psi}) & d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\Delta\boldsymbol{s}) \end{Bmatrix}. \tag{66}$$

Finally, the solution of the above problem requires the solution of the linear system

$$\begin{Bmatrix} k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{u}) & p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\Delta\boldsymbol{s}) \\ p(\boldsymbol{u},\boldsymbol{s};\Delta\boldsymbol{u},\boldsymbol{\psi}) & d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\Delta\boldsymbol{s}) \end{Bmatrix} = -\begin{Bmatrix} \mathcal{R}(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta}) \\ \mathcal{G}(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi}) \end{Bmatrix} \tag{67}$$

in each Newton step.

Remark 11 The diagonal elements of the above saddle point problem are the pseudo load operators of the physical (43) and material residual (51). Therefore, the physical and the material problem are coupled by the pseudo load operator $p(\cdot,\cdot)$, which is used for the solution of structural optimization problems. Therefore, these quantities are already calculated and we have to calculate additionally only the material stiffness $d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\Delta\boldsymbol{s})$ in order to solve this problem with standard nonlinear programming algorithms.

An explicit formulation of all operators for a simple linear model problem is given in Appendix A and for nonlinear hyperelasticity in [36].

4 The FE approximation and solution algorithms

4.1 The discrete optimization problem

The finite element formulation is based on a conforming Galerkin method defined on meshes $\mathcal{T}_h = \{K\}$ with a mesh parameter h, consisting of closed cells K which are either triangles or quadrilaterals. The boundary ∂K of each element K is assumed to be Lipschitz-continuous. On the mesh $\mathcal{T}_h$ we define finite dimensional element spaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{D}_h \subset \mathcal{D}$ consisting of cellwise polynomial functions.

The corresponding discrete state $\boldsymbol{u}_h \in \mathcal{V}_h \subset \mathcal{V}$ and discrete design $\boldsymbol{s}_h \in \mathcal{D}_h \subset \mathcal{D}$ are determined by the following discrete version of Problem 2.

Problem 6 Find $\{\boldsymbol{u}_h, \boldsymbol{s}_h\} \in \mathcal{V}_h \times \mathcal{D}_h$ of the objective function $\mathcal{J} : \mathcal{V}_h \times \mathcal{D}_h \to \mathbb{R}$ such that

$$\mathcal{J}(\boldsymbol{u}_h, \boldsymbol{s}_h) = E(\boldsymbol{u}_h, \boldsymbol{s}_h) \to \min_{\{\boldsymbol{u}_h, \boldsymbol{s}_h\} \in \mathcal{V}_h \times \mathcal{D}_h} \tag{68}$$

subject to the constraint

$$\langle \boldsymbol{A}(\boldsymbol{u}_h, \boldsymbol{s}_h), \boldsymbol{\eta}_h \rangle = \mathcal{R}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h) = 0 \quad \forall\, \boldsymbol{\eta}_h \in \mathcal{V}_h. \tag{69}$$

For the optimal solutions $\boldsymbol{u}_h^* \in \mathcal{V}_h$ and $\boldsymbol{s}_h^* \in \mathcal{D}_h$ in the chosen approximation spaces $\mathcal{V}_h$ and $\mathcal{D}_h$ we require that

$$\mathcal{J}(\boldsymbol{u}_h^*, \boldsymbol{s}_h^*) \leq \mathcal{J}(\boldsymbol{u}_h, \boldsymbol{s}_h) \quad \forall\, \boldsymbol{u}_h, \boldsymbol{s}_h \in \mathcal{V}_h \times \mathcal{D}_h. \tag{70}$$

The error of the state $\boldsymbol{e}_{u,h}$ and the design $\boldsymbol{e}_{s,h}$ in the chosen approximation spaces are introduced by

$$\boldsymbol{e}_{u,h} := \boldsymbol{u}_h^* - \boldsymbol{u}_h \quad \text{and} \quad \boldsymbol{e}_{s,h} := \boldsymbol{s}_h^* - \boldsymbol{s}_h \tag{71}$$

respectively.

4.2 The error in the material residual

The fulfillment of the physical residual for every admissible design is an important constraint of the optimization problem, i.e.

$$\mathcal{R}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h) = 0 \quad \forall\, \boldsymbol{s}_h \in \mathcal{D}_h. \tag{72}$$

The optimality condition (28) for the material problem holds only for stationary points $\boldsymbol{u}_h^*$ and $\boldsymbol{s}_h^*$ of the Lagrangian, i.e. $\mathcal{L}_s' = \mathcal{G}(\boldsymbol{u}_h^*, \boldsymbol{s}_h^*; \boldsymbol{\psi}_h) = 0$. This means, that the material residual is not fulfilled for every $\boldsymbol{u}_h \neq \boldsymbol{u}_h^* = \boldsymbol{u}_h(\boldsymbol{s}_h^*)$, i.e.

$$\mathcal{G}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\psi}_h) \neq 0 \quad \forall\, \boldsymbol{u}_h \neq \boldsymbol{u}_h^* = \boldsymbol{u}_h(\boldsymbol{s}_h^*) \in \mathcal{V}_h. \tag{73}$$

Hence, we obtain a material residual as a result of the non-optimal solution $\boldsymbol{u}_h$, which is a result of the non-optimal design $\boldsymbol{s}_h$ in the sense of the minimization of Problem 6. With the definition of the errors (71) we have $\boldsymbol{s}_h^* = \boldsymbol{s}_h + \boldsymbol{e}_{s,h}$ and therefore

$$\mathcal{G}(\boldsymbol{u}_h^*, \boldsymbol{s}_h^*; \boldsymbol{\psi}_h) = \mathcal{G}(\boldsymbol{u}_h(\boldsymbol{s}_h + \boldsymbol{e}_{s,h}), \boldsymbol{s}_h + \boldsymbol{e}_{s,h}; \boldsymbol{\psi}_h) = 0. \tag{74}$$

From this, a suitable approximation for the error could be obtained from the linearization

$$\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h) + D_s\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h) \cdot \boldsymbol{e}_{s,h} + \mathcal{O} = 0. \tag{75}$$

Finally, with the tangent operator of the material problem

$$\begin{aligned} &b_T(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h, \boldsymbol{e}_{s,h}) \\ &\quad := D_s\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h) \cdot \boldsymbol{e}_{s,h} \\ &\quad = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\mathcal{G}(\hat{\boldsymbol{u}}_h, \boldsymbol{s}_h + \varepsilon \boldsymbol{e}_{s,h}; \boldsymbol{\psi}_h)\right]\Big|_{\varepsilon=0} \\ &\qquad + \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\mathcal{G}(\boldsymbol{u}_h(\boldsymbol{s}_h + \varepsilon \boldsymbol{e}_{s,h}), \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h)\right]\Big|_{\varepsilon=0} \end{aligned} \tag{76}$$

and the disregard of the remainder $\mathcal{O}$ an equation for the error in the design or rather for the error in the material residual follows in the form

$$b_T(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h, \boldsymbol{e}_{s,h}) = -\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h) \quad \forall\, \boldsymbol{\psi}_h \in \mathcal{D}_h. \tag{77}$$

This is used for a staggered solution algorithm, see Sect. 4.5.2. Furthermore, it could be used for a mesh optimization algorithm, see Sect. 5.

4.3 The discrete full Newton algorithm

The linearization of the Lagrangian yields the finite element approximation of (67) in the form

$$\begin{Bmatrix} k(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h, \Delta\boldsymbol{u}_h) & p(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h, \Delta\boldsymbol{s}_h) \\ p(\boldsymbol{u}_h, \boldsymbol{s}_h; \Delta\boldsymbol{u}_h, \boldsymbol{\psi}_h) & d(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\psi}_h, \Delta\boldsymbol{s}_h) \end{Bmatrix} = - \begin{Bmatrix} \mathcal{R}_h \\ \mathcal{G}_h \end{Bmatrix} \tag{78}$$

where $\mathcal{R}_h = \mathcal{R}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h)$ and $\mathcal{G}_h = \mathcal{G}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\psi}_h)$. For a matrix description we introduce the discrete approximations for displacements and design, i.e. the nodal displacement vector $\boldsymbol{u} \in \mathbb{R}^n$ and the vector of design variables $\boldsymbol{s} \in \mathbb{R}^m$. Here, n and m are the dimensions of the introduced approximation spaces.

After a standard finite element discretization, the system (78) takes the form

$$\begin{bmatrix} \boldsymbol{K} & \boldsymbol{P} \\ \boldsymbol{P}^{\mathrm{T}} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{u} \\ \Delta\boldsymbol{s} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{G} \end{bmatrix}, \tag{79}$$

where $\boldsymbol{K} \in \mathbb{R}^{n \times n}$ is the usual tangent stiffness matrix and $\boldsymbol{P} \in \mathbb{R}^{n \times m}$ is the tangent pseudo load operator matrix as well as $\boldsymbol{D} \in \mathbb{R}^{m \times m}$ is the tangent material stiffness matrix associated to the bilinear forms $k(\cdot,\cdot)$, $p(\cdot,\cdot)$ and $d(\cdot,\cdot)$. Furthermore, $\boldsymbol{R} \in \mathbb{R}^n$ denotes the physical residual vector and $\boldsymbol{G} \in \mathbb{R}^m$ the material residual vector associated to the functionals $\mathcal{R}(\boldsymbol{u}_h, \boldsymbol{s}_h; \cdot)$ and $\mathcal{G}(\boldsymbol{u}_h, \boldsymbol{s}_h; \cdot)$, respectively.

Remark 12 We solve simultaneously the physical and material residual problem. The system (79) has a typical saddle point structure and we have to be careful by solving this system. Often a preconditioner is needed in order to obtain a

well-conditioned system matrix. See for instance [11] and the references therein for preconditioners for this type of equations.

Remark 13 If we substitute the first equation of (79) into the second and vice versa, we can eliminate $\Delta \boldsymbol{u}$ or $\Delta \boldsymbol{s}$ by using the Schur complement in order to solve the problem one after the other and to minimize the size of the system. We obtain a formulation for the state variables

$$[\boldsymbol{K} - \boldsymbol{P}\boldsymbol{D}^{-1}\boldsymbol{P}^{\mathrm{T}}]\Delta \boldsymbol{u} = \boldsymbol{P}\boldsymbol{D}^{-1}\boldsymbol{G} - \boldsymbol{R} \tag{80}$$

or for the design variables

$$[\boldsymbol{D} - \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{P}]\Delta \boldsymbol{s} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{R} - \boldsymbol{G}. \tag{81}$$

This requires well-conditioned matrices $\boldsymbol{K}$ and $\boldsymbol{D}$ in order to compute the corresponding inverse matrices accurately and to obtain a stable solution algorithm.

4.4 The discrete sensitivity equations

With the above definitions, the discrete versions of the sensitivity equations for the physical (44) and material (55) residuals become

$$\boldsymbol{K}\delta \boldsymbol{u} + \boldsymbol{P}\delta \boldsymbol{s} = \boldsymbol{0} \quad \text{or} \quad \delta \boldsymbol{u} = -\boldsymbol{K}^{-1}\boldsymbol{P}\delta \boldsymbol{s} \tag{82}$$

and

$$\boldsymbol{P}^{\mathrm{T}}\delta \boldsymbol{u} + \boldsymbol{D}\delta \boldsymbol{s} = \boldsymbol{0} \quad \text{or} \quad \delta \boldsymbol{s} = -\boldsymbol{D}^{-1}\boldsymbol{P}^{\mathrm{T}}\delta \boldsymbol{u}. \tag{83}$$

For chosen fixed variations $\delta \hat{\boldsymbol{s}}$ and $\delta \hat{\boldsymbol{u}}$ the discrete versions of the sensitivity equations for the physical (45) and material (56) problem follow in the form

$$\boldsymbol{K}\delta \boldsymbol{u} = -\boldsymbol{Q}_p \quad \text{with} \quad \boldsymbol{Q}_p := \boldsymbol{P}\delta \hat{\boldsymbol{s}} \tag{84}$$

and

$$\boldsymbol{D}\delta \boldsymbol{s} = -\boldsymbol{Q}_m \quad \text{with} \quad \boldsymbol{Q}_m := \boldsymbol{P}^{\mathrm{T}}\delta \hat{\boldsymbol{u}} \tag{85}$$

respectively. Here, $\boldsymbol{Q}_p \in \mathbb{R}^n$ is the pseudo load vector of the physical residual problem associated to the functional $Q(\boldsymbol{u}_h, \boldsymbol{s}_h; \cdot, \delta \hat{\boldsymbol{s}}_h)$ and $\boldsymbol{Q}_m \in \mathbb{R}^m$ is the pseudo load vector of the material residual problem associated to the functional $Q(\boldsymbol{u}_h, \boldsymbol{s}_h; \delta \hat{\boldsymbol{u}}_h, \cdot)$.

Remark 14 It is important to note, that we obtain with the relations from (82) directly a connection between the physical and the material space. Let $\boldsymbol{S}_p := -\boldsymbol{K}^{-1}\boldsymbol{P} \in \mathbb{R}^{n\times m}$ be the discrete sensitivity operator, the physical and material space are connected by the transformation

$$\delta \boldsymbol{u} = \boldsymbol{S}_p\, \delta \boldsymbol{s}. \tag{86}$$

We call the matrix $\boldsymbol{S}_p$ the *sensitivity matrix* of the physical problem. With the knowledge of the pseudo load operator matrix $\boldsymbol{P}$, we can evaluate the sensitivity equation for arbitrary admissible variations $\delta \hat{\boldsymbol{s}}$ in the material space. In the same manner, the discrete sensitivity operator of the material problem, i.e. the *sensitivity matrix* $\boldsymbol{S}_m := -\boldsymbol{D}^{-1}\boldsymbol{P}^{\mathrm{T}} \in \mathbb{R}^{m\times n}$ follows from (83) and we obtain the transformation

$$\delta \boldsymbol{s} = \boldsymbol{S}_m\, \delta \boldsymbol{u}. \tag{87}$$

With this, we can perform the sensitivity analysis for arbitrary admissible variations $\delta \hat{\boldsymbol{u}}$ in the physical space.

Remark 15 In Remark 10 we have mentioned the sensitivity of the energy release rate $G' = -\mathcal{G}' = -(d(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}, \delta \boldsymbol{s}) + p(\boldsymbol{u}, \boldsymbol{s}; \delta \boldsymbol{u}, \boldsymbol{\psi}))$. With the notations from (82) and (83) the discrete approximation follows in the form

$$G'_h = -\mathcal{G}'(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\psi}_h, \delta \boldsymbol{u}_h, \delta \boldsymbol{s}_h) = -\boldsymbol{\psi}^{\mathrm{T}}\delta \boldsymbol{G} \tag{88}$$

where

$$\begin{aligned}\delta \boldsymbol{G} &= \boldsymbol{D}\delta \boldsymbol{s} + \boldsymbol{P}^{\mathrm{T}}\delta \boldsymbol{u} = \boldsymbol{D}\delta \boldsymbol{s} - \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{P}\, \delta \boldsymbol{s} \\ &= [\boldsymbol{D} - \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{P}]\, \delta \boldsymbol{s}\end{aligned} \tag{89}$$

is the variation of the material residual. With the sensitivity operator $\boldsymbol{S}_p$ of the physical problem (86) we obtain

$$\delta \boldsymbol{G} = \left[\boldsymbol{D} + \boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_p\right]\delta \boldsymbol{s}. \tag{90}$$

Hence, for a given variation in the design $\delta \boldsymbol{s}$ we can calculate the variation in the energy release rate or rather the variation in the material or configurational residual $\boldsymbol{G}$.

4.5 Alternative solution algorithms

Alternatively, other solution algorithms could be used in order to solve the optimization problem.

4.5.1 Steepest descent method

The simplest way to solve the optimization problem is the well-known steepest descent method. We rewrite the objective function to $\mathcal{J}(\boldsymbol{s}) = \mathcal{J}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s})$.

For a given starting point $\boldsymbol{s}_0$ and $\boldsymbol{u}(\boldsymbol{s}_0)$ we calculate a search direction $\boldsymbol{v} := -\mathcal{J}'_s$ and perform a line search with $\varepsilon > 0$ such that $\mathcal{J}(\boldsymbol{s} + \varepsilon \boldsymbol{v}) < \mathcal{J}(\boldsymbol{s})$. This leads to the solution algorithm

$$\boldsymbol{s}_{i+1} = \boldsymbol{s}_i + \varepsilon_i\, \boldsymbol{v}_i = \boldsymbol{s}_i - \varepsilon_i\, \mathcal{J}'_s(\boldsymbol{u}, \boldsymbol{s}_i). \tag{91}$$

In the case, that we choose the energy of the system (25) as the objective functional, i.e. $\mathcal{J} = E$, the discrete update scheme is given by

$$\boldsymbol{s}_{i+1} = \boldsymbol{s}_i + \Delta \boldsymbol{s} = \boldsymbol{s}_i - \varepsilon_i\, \boldsymbol{G}(\boldsymbol{u}_h, \boldsymbol{s}_{h,i}) \tag{92}$$

where $\boldsymbol{G}(\boldsymbol{u}_h, \boldsymbol{s}_{h,i}) \in \mathbb{R}^m$ is the material residual vector associated to the functional $\mathcal{G}(\boldsymbol{u}_h, \boldsymbol{s}_h; \cdot)$ at the current state $\boldsymbol{u}_h(\boldsymbol{s}_{h,i})$.

Remark 16 Additional we have to fulfil in every step the constraint $\boldsymbol{A}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{v}_h) = 0$, which means that we have to solve the equilibrium equation for the primal problem. Moreover, the computation of the accurate material residual $\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h)$ for the next iteration step requires the solution of the state variable $\boldsymbol{u}_h$ for the updated current design, i.e. we have to solve $\mathcal{R}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\eta}_h) = 0$. For a simple linear physical problem, only the solution of one linear equation is required. In the general nonlinear case, we have to perform some Newton iterations for the physical problem in order to find the new state of equilibrium for the current design, i.e. we have to solve the tangent version of the primal problem

$$k(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\eta}_h, \Delta\boldsymbol{u}_h) = -\mathcal{R}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\eta}_h) \tag{93}$$

in every Newton step. The best performance of the algorithm could be achieved, if we transfer beforehand the data of the state $\boldsymbol{u}_h$ to the new design. To do this, we can use the sensitivity relation for the physical problem (86) and obtain with a sufficient small $\Delta\boldsymbol{s}$ an approximation for the change in the state from

$$\Delta\boldsymbol{u} = -\boldsymbol{K}^{-1}\boldsymbol{P}\,\Delta\boldsymbol{s} = \boldsymbol{S}_p\,\Delta\boldsymbol{s}. \tag{94}$$

After this, the system is usually still unbalanced, but only few Newton iterations are required in order to find the new state of equilibrium.

Remark 17 The steepest descent directions at two consecutive steps could be orthogonal to each other. This tends to slow down the steepest descent method although it is convergent and hence a large number of iterations are required in order to find the solution. A simple and efficient improvement is the *conjugate gradient method*. This method is also based only on gradient information of the objective functional, see e.g. [45] for more details. The conjugate gradient method should be preferred over the steepest descent method, because the rate of convergence could be improved without appreciable cost.

4.5.2 Staggered solution method

Alternatively, we propose a staggered solution method, i.e. we solve the physical problem with a fixed design $\hat{\boldsymbol{s}}$ and thereafter the material problem with a fixed deformation $\hat{\boldsymbol{u}}$.

For a given solution $\hat{\boldsymbol{u}}$ of the physical residual $\mathcal{R}$, we have to solve $\mathcal{G}(\hat{\boldsymbol{u}}, \boldsymbol{s}; \boldsymbol{\psi}) = 0$ within a Newton algorithm. This requires the total partial derivative (17) of the functional $\mathcal{G}(\boldsymbol{u}, \boldsymbol{s}) = \mathcal{G}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s})$ with respect to $\boldsymbol{s}$, which is given by

$$D_s\mathcal{G}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s}) = \frac{\partial\mathcal{G}}{\partial\boldsymbol{s}} + \frac{\partial\mathcal{G}}{\partial\boldsymbol{u}}\,\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{s}}. \tag{95}$$

With this, the linearization of $\mathcal{G}$ in the direction $\Delta\boldsymbol{s}$ yields

$$\mathcal{G}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}, \boldsymbol{\psi}) + D_s\mathcal{G}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}, \boldsymbol{\psi}) \cdot \Delta\boldsymbol{s} + \mathcal{O} = 0 \tag{96}$$

with the tangent operator

$$\begin{aligned} &b_T(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s}; \boldsymbol{\psi}, \Delta\boldsymbol{s}) \\ &\quad := D_s\mathcal{G}(\boldsymbol{u}, \boldsymbol{s}, \boldsymbol{\psi}) \cdot \Delta\boldsymbol{s} \\ &\quad = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\mathcal{G}(\boldsymbol{u}, \boldsymbol{s} + \varepsilon\Delta\boldsymbol{s}; \boldsymbol{\psi}) + \mathcal{G}(\boldsymbol{u}(\boldsymbol{s} + \varepsilon\Delta\boldsymbol{s}), \boldsymbol{s}; \boldsymbol{\psi})\right]\Big|_{\varepsilon=0}. \end{aligned} \tag{97}$$

The remainder $\mathcal{O}$ can usually be neglected and we obtain a linear equation for $\Delta\boldsymbol{s}$ in the form

$$b_T(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}; \boldsymbol{\psi}, \Delta\boldsymbol{s}) = -\mathcal{G}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{s}}, \boldsymbol{\psi}) \quad \forall\,\boldsymbol{\psi} \in \mathcal{D}. \tag{98}$$

The finite element approximation reads

$$b_T(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h, \Delta\boldsymbol{s}_h) = -\mathcal{G}(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h, \boldsymbol{\psi}_h) \quad \forall\,\boldsymbol{\psi}_h \in \mathcal{D}_h. \tag{99}$$

This is equivalent to Eq. (77), in which we have replaced the error in the design $\boldsymbol{e}_{s,h}$ by the design increment $\Delta\boldsymbol{s}_h$.

Remark 18 In Sect. 4.2 we have mentioned, that the material residual is not fulfilled for every $\boldsymbol{u}_h \neq \boldsymbol{u}_h^* = \boldsymbol{u}_h(\boldsymbol{s}_h^*)$, i.e. $\mathcal{G}(\boldsymbol{u}_h, \boldsymbol{s}_h; \boldsymbol{\psi}_h) \neq 0\,,\ \forall\,\boldsymbol{u}_h \neq \boldsymbol{u}_h^* \in V_h$ and hence, for a given solution $\boldsymbol{u}_h$ of the physical problem, we have to solve the material problem with a Newton algorithm in order to find a $\boldsymbol{s}_h^*$ such that $\boldsymbol{u}_h(\boldsymbol{s}_h^*) = \boldsymbol{u}_h^*$ and hence $\mathcal{G}(\boldsymbol{u}_h^*, \boldsymbol{s}_h^*; \boldsymbol{\psi}_h) = 0$.

In the discrete case follows from Eq. (51) $\partial_u\boldsymbol{G} = \boldsymbol{P}^{\mathrm{T}}$ and from Eq. (82) $\partial_s\boldsymbol{u} = -\boldsymbol{K}^{-1}\boldsymbol{P}$. Hence, a matrix description is given by

$$\begin{aligned} b_T(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{s}}_h; \boldsymbol{\psi}_h, \Delta\boldsymbol{s}_h) &= \boldsymbol{\psi}^{\mathrm{T}}\left[\frac{\partial\boldsymbol{G}}{\partial\boldsymbol{s}} + \frac{\partial\boldsymbol{G}}{\partial\boldsymbol{u}}\,\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{s}}\right]\Delta\boldsymbol{s} \\ &= \boldsymbol{\psi}^{\mathrm{T}}\left[\boldsymbol{D} - \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{P}\right]\Delta\boldsymbol{s}. \end{aligned} \tag{100}$$

Finally, we have to solve the discrete version of (98), i.e.

$$\left[\boldsymbol{D} - \boldsymbol{P}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{P}\right]\Delta\boldsymbol{s} = \left[\boldsymbol{D} + \boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_p\right]\Delta\boldsymbol{s} = -\boldsymbol{G} \tag{101}$$

in every Newton step. Here, $\boldsymbol{S}_p = -\boldsymbol{K}^{-1}\boldsymbol{P}$ is the sensitivity operator of the physical problem (86).

In order to update the state variables $\boldsymbol{u}_h$ during the Newton iteration, we can use the same procedure as mentioned in Remark 16.

Remark 19 In a classical staggered solution algorithm, the coupling terms $\boldsymbol{P}$ in (79) are usually simply neglected. As a result of the functional dependencies $\mathcal{G}(\boldsymbol{u}, \boldsymbol{s}) = \mathcal{G}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s})$ of the material residual, we have nevertheless the coupling terms in the tangent operator (97) for the solution of the material motion problem.

5 Application to mesh optimization

In an abstract sense, mesh optimization is a structural optimization problem, see for instance [15,16,40,49] for first steps in this direction based on a discrete formulation.

In the case of a mesh optimization problem the nodal coordinates $\boldsymbol{X}$ are chosen as design variables, i.e. $\boldsymbol{s} = \boldsymbol{X}$. The optimal nodal positions (the design) of a given mesh (the reference configuration) can be computed by solving Eq. (79), where the design variables $\boldsymbol{s}$ are replaced by the vector of nodal coordinates $\boldsymbol{X}$, i.e.

$$\begin{bmatrix} \boldsymbol{K} & \boldsymbol{P} \\ \boldsymbol{P}^{\mathrm{T}} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{u} \\ \Delta \boldsymbol{X} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{G} \end{bmatrix}. \tag{102}$$

All inner nodes or a subset of nodes are chosen as design variables. Additional we can include nodal coordinates in tangential direction on the boundary. Finally, it follows a r-adaption algorithm for mesh optimization [33,34].

Alternatively, other solution algorithms could be used for the mesh optimization problem. The simple steepest descent method (92) yields an algorithm for mesh optimization in the form

$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i - \varepsilon\, \boldsymbol{G}(\boldsymbol{u}_h, \boldsymbol{X}_i). \tag{103}$$

In the context of configurational mechanics and mesh optimization, such an algorithm was used for instance by [42,43].

In Sect. 4.5.2 we have proposed a staggered solution algorithm. For this, the linearized material problem (101) yields the form

$$\left[\boldsymbol{D} + \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_p\right] \Delta \boldsymbol{X} = -\boldsymbol{G}. \tag{104}$$

A similar expression for the full Newton algorithm (102) was obtained by [50,51] and [1,31]. The latter authors have also proposed a staggered solution method, which disregards the coupling terms $\boldsymbol{P}$ in (102), see Remark 19.

Remark 20 The material residual as an error indicator for mesh optimization could also be used as indicator within a h-adaptivity algorithm, see e.g. [44]. Furthermore, it seems promising to combine r- and h-adaptivity as well as remeshing strategies based on the above mentioned error indicators. The development of practical error estimators for the discretization error is desirable.

5.1 Numerical problems and regularization

The mesh optimization problem can be interpreted as an inverse problem. It requires the solution of the material residual problem within a staggered algorithm (104) or the coupled system (102). Different numerical difficulties arise due to the solution of this problem. In many cases, the Hessian matrix of the system is ill-conditioned and becomes singular or close to singular during the iterations and therefore the Newton algorithm is not stable and fails. The optimization problem is non-convex in general and hence, the solution needs not to be unique. Therefore, the problem could be termed ill-posed in the sense of Hadamard and reasonable regularization methods should be used in order to regularize the problem. A classical regularization is given by adding a penalty functional $\mathcal{P}$ to the original objective $\mathcal{J}$, i.e.

$$\tilde{\mathcal{J}}(\boldsymbol{u}, \boldsymbol{s}) = \mathcal{J}(\boldsymbol{u}, \boldsymbol{s}) + \mathcal{P}(\gamma, \boldsymbol{s}) \tag{105}$$

where γ is a penalty parameter. The choice of the penalty functional and the penalty parameter depends on the problem. For a classical Tikhonov-type regularization we have often a functional of the form

$$\mathcal{P}(\gamma, \boldsymbol{s}) = \frac{\gamma}{2} ||\boldsymbol{s} - \boldsymbol{s}_0||^2, \tag{106}$$

see for instance [20] and the references therein for details. The numerical difficulties for the mesh optimization problem are also mentioned by [1,41]. The authors have proposed different strategies in order to overcome these problems.

An ill-conditioned problem is indicated by a large condition number c of the system matrix $\boldsymbol{A}$ from the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, which can be defined by

$$c(\boldsymbol{A}) = \frac{\sigma_{\max}(\boldsymbol{A})}{\sigma_{\min}(\boldsymbol{A})}. \tag{107}$$

Here, $\sigma_{\max}(\boldsymbol{A})$ and $\sigma_{\min}(\boldsymbol{A})$ are the maximal and minimal *singular values* of $\boldsymbol{A}$, respectively. A large condition number is caused by nearly zero singular values of $\boldsymbol{A}$. In the context of sensitivity analysis, the condition number quantifies the sensitivity of the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with respect to small perturbations $(\boldsymbol{A} + \varepsilon \tilde{\boldsymbol{A}})\boldsymbol{x}(\varepsilon) = \boldsymbol{b} + \varepsilon \tilde{\boldsymbol{b}}$. Small changes in $\boldsymbol{A}$ or $\boldsymbol{b}$ can induce large changes in $\boldsymbol{x}$ if the condition number in large. The singular values can be obtained by a *singular value decomposition* (SVD) of the system matrix. The SVD is used in many fields of engineering and physics, which deal with inverse problems.

For a rectangular matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, the SVD is a decomposition of the form

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{Y}^{\mathrm{T}} = \sum_{i=1}^{m} \sigma_i \boldsymbol{v}_i \boldsymbol{y}_i^{\mathrm{T}} \tag{108}$$

where $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{Y} \in \mathbb{R}^{m \times m}$ are matrices with orthonormal columns, i.e. $\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{I} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}} = \boldsymbol{I} \in \mathbb{R}^{m \times m}$. The matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_m) \in \mathbb{R}^{n \times m}$ is a diagonal matrix with non-negative diagonal elements σ_i. The quantities σ_i are called singular values of $\boldsymbol{A}$ and they appearing in decreasing order such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$. The column vectors of $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ and $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m]$ are called *left and right singular vectors* of $\boldsymbol{A}$, respectively, see e.g. [23] for details.

We use the singular value decomposition in order to calculate the condition number and therefore information about the solution behavior as well as for a regularization of the Hessian matrix.

The SVD is a reliable method in order to identify the numerical rank. The singular vectors, which are corresponding to the non-zero singular values σ_i of $\boldsymbol{A}$ span the range

of $\boldsymbol{A}$. With these, we can define the numerical rank r of $\boldsymbol{A}$ as the number of the non-zero singular values of $\boldsymbol{A}$, which are larger than a defined tolerance. The solution of the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ could be expressed in terms of the left and right singular vectors

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \boldsymbol{Y}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{b} = \sum_{i=1}^{m} \frac{1}{\sigma_i}\, \boldsymbol{y}_i\, \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{b}. \tag{109}$$

The solution $\boldsymbol{x}$ is given by linear combination of $\boldsymbol{y}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{b}$ weighted with the inverse singular values σ_i^{-1}. At this point, we see the influence of nearly zero singular values σ_i, i.e. we have

$$\lim_{\sigma_i \to 0} \frac{1}{\sigma_i}\, \boldsymbol{y}_i\, \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{b} = \infty. \tag{110}$$

If we truncate the sum in Eq. (109) after r terms, the result is a rank-r approximation to the original Hessian. The error in the approximation depends upon the magnitude of the neglected singular values. This could be interpreted as filtering out the noisy data.

Furthermore, the singular value decomposition provide information about the most relevant design variables during the optimization. In the context of mesh optimization, the SVD of the sensitivity relation for the material residual (90)

$$\delta\boldsymbol{G} = \left[\boldsymbol{D} + \boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_p\right]\delta\boldsymbol{X} \tag{111}$$

could be used in order to find out the most relevant nodal coordinates (design variables). Nodal coordinates, which do not contribute to a change in the energy during the mesh optimization process should not be considered in order to reduce the size of the Hessian and to avoid noisy data. The regularization and the identification of the most relevant design variables are discussed in more detail in [35].

5.2 Mesh distortion control

The quality of the mesh (the distortion of the mesh) has an important influence on the shape derivatives and hence on the results and the success of the optimization process.

In order to control the mesh distortion we use a simple geometrical distortion parameter ξ, which is used in shape optimization [54]. The distortion parameter for a quadrilateral element with four nodes is given by

$$\xi = \frac{4}{A}\, \min(\det \boldsymbol{J}_i) \quad i = 1, \ldots, 4. \tag{112}$$

Here, $\min(\det \boldsymbol{J}_i)$ denotes the minimum value of the Jacobian determinant and A the element area. The distortion parameter has the following properties:

$$\xi \begin{cases} = 1 & \text{element is a parallelogram} \\ > 0 & \text{element is convex} \\ = 0 & \text{element is degenerate (triangle)} \\ < 0 & \text{element is concave} \end{cases}$$

The parameter ξ must be greater than zero to avoid degeneracy of the element. In practical computations, the condition $\xi \geq \mathrm{tol}_\xi$ is used with a tolerance $\mathrm{tol}_\xi \in [0.1, 0.2]$.

5.3 Overall solution algorithm

We consider a staggered solution algorithm for Eq. (104). The overall solution scheme for mesh optimization is given as follows.

1. Solve the physical problem within a Newton Method.
2. Solve the material problem: Compute the material tangent operator $\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_p$. Check the condition number and the numerical rank of $\boldsymbol{A}$ with SVD. If necessary, compute a rank-r approximation in order to overcome the ill-conditioning. Compute $\Delta\boldsymbol{X}$.
3. Line search: Compute α such that $\xi(\alpha) \geq \mathrm{tol}_\xi$. Compute β such that $\mathcal{J}(\boldsymbol{X} + \beta\Delta\boldsymbol{X}) < \mathcal{J}(\boldsymbol{X})$.
4. Update $\boldsymbol{X}$ with $\Delta\boldsymbol{X} = \varepsilon\, \Delta\boldsymbol{X}$, where $\varepsilon = \min[\alpha, \beta]$.
5. Update $\boldsymbol{u}$ with sensitivity relation $\Delta\boldsymbol{u} = \boldsymbol{S}_p\Delta\boldsymbol{X}$. After that, the system is usually still unbalanced, see Remark 16. Goto 1 and find the new state of equilibrium.

The line search consists of two parts. In the first part, we compute a step size parameter α in order to avoid mesh distortion. For a given tolerance tol_ξ, the step size parameter is given explicitly from (112)

$$\xi(\alpha) = \frac{4}{A(\boldsymbol{X} + \alpha\Delta\boldsymbol{X})} \det \boldsymbol{J}(\boldsymbol{X} + \alpha\Delta\boldsymbol{X}) \geq \mathrm{tol}_\xi. \tag{113}$$

The maximum step length α along a given search direction can be computed in closed form from this expression [54]. The second part is the computation of the usual step size β used in nonlinear programming in order to guarantee the decrease in the energy. For this, we use the Armijo-Goldstein condition.

An algorithm for the simultaneous solution of the physical and material problem (102) could be obtained in the same manner by using the regularization and line search as in the staggered solution algorithm.

5.4 Example 1: unit square

We consider a two-dimensional unit square $\Omega = (0, 1)^2$ with homogeneous Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ on Γ. The body is loaded with a horizontal body load $b_x = 50$. The considered model problem is given in Appendix A. The

problem is modeled with the plane strain condition with $E = 1,000$ and $\nu = 0.3$. We discretize the model with 153 Q4 elements and we choose an arbitrary irregular initial discretization, see Fig. 1a. The optimal solution $\boldsymbol{X}_h^*$ (the optimal design) for $\mathcal{J}(\boldsymbol{u}_h, \boldsymbol{X}_h)$ for this simple model problem under the given load case is a symmetric discretization as shown in Fig. 1b. This discretization leads to the best approximation for the given problem in the chosen approximation spaces $\mathcal{V}_h$ and $\mathcal{D}_h$. The mesh is finer in the middle of the domain and tends to a coarser discretization on the boundary. This behavior reflects the solution for the given load case. The solution $\boldsymbol{u}_h$ has the maximum in the middle and goes to zero on the boundary, see Fig. 1c.

We consider the simultaneous solution of the physical and material problem. In order to find the minimum of the objective $\mathcal{J}(\boldsymbol{u}, \boldsymbol{X})$ we have to solve Eq. (102) in each Newton step. The mesh distortion control parameter (112) with $\text{tol}_\xi = 0.2$ was used. The results of the full Newton algorithm are given in Table 1. The solution $\mathcal{J}(\boldsymbol{u}_h^*, \boldsymbol{X}_h^*)$ was obtained after seven iterations, see Fig. 1b. For simplicity, we allow in this first example only horizontal movements of the inner nodes. During the iteration the free nodes move in the opposite direction of the material residual forces and the energy decrease from $\mathcal{J}(\boldsymbol{u}_h^0, \boldsymbol{X}_h^0) = -5.2323 \times 10^{-2}$ (initial mesh) to $\mathcal{J}(\boldsymbol{u}_h^*, \boldsymbol{X}_h^*) = -5.2466 \times 10^{-2}$ (optimized mesh), see Table 1. Here, $\boldsymbol{u}_h^* \in \mathcal{V}_h$ and $\boldsymbol{X}_h^* \in \mathcal{D}_h$ are the optimal solution in the chosen approximation spaces $\mathcal{V}_h$ and $\mathcal{D}_h$ in the sense of the minimization problem (6). With the optimal solutions $\{\boldsymbol{u}_h^*, \boldsymbol{X}_h^*\}$, we introduce the error in the energy $e_{\mathcal{J},h} := \mathcal{J}(\boldsymbol{u}_h^*, \boldsymbol{X}_h^*) - \mathcal{J}(\boldsymbol{u}_h, \boldsymbol{X}_h)$, the error in the state $\boldsymbol{e}_{u,h} := \boldsymbol{u}_h^* - \boldsymbol{u}_h$ and the design $\boldsymbol{e}_{X,h} := \boldsymbol{X}_h^* - \boldsymbol{X}_h$, respectively, see Eq. (71). The relative errors are given by

$$\eta_{\mathcal{J}}^{\text{rel}} := \left| \frac{e_{h,\mathcal{J}}}{\mathcal{J}(\boldsymbol{u}_h^*, \boldsymbol{X}_h^*)} \right| \tag{114}$$

as well as

$$\eta_u^{\text{rel}} := \frac{||\boldsymbol{e}_{h,u}||_{L_2}}{||\boldsymbol{u}_h^*||_{L_2}} \quad \text{and} \quad \eta_X^{\text{rel}} := \frac{||\boldsymbol{e}_{h,X}||_{L_2}}{||\boldsymbol{X}_h^*||_{L_2}}, \tag{115}$$

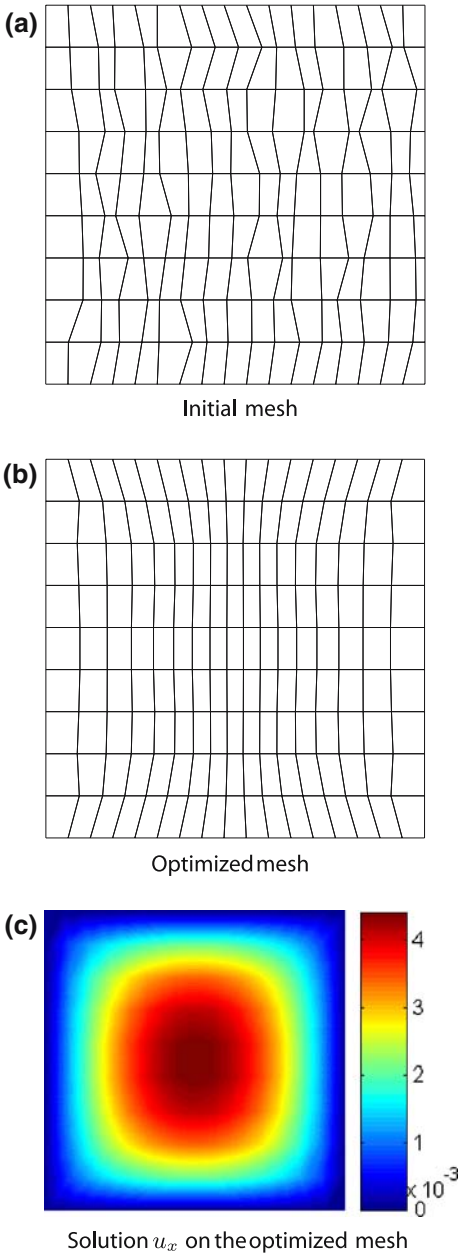


Fig. 1 Unit square with horizontal body load

Table 1 Results of the full Newton algorithm

i	$\|\|\boldsymbol{R}\|\|$	$\|\|\boldsymbol{G}\|\|$	$\|\|\varepsilon\,\Delta\boldsymbol{u}\|\|$	$\|\|\varepsilon\,\Delta\boldsymbol{X}\|\|$	$\mathcal{J}(\boldsymbol{u}_h, \boldsymbol{X}_h)$	$\varepsilon = \min[\alpha, \beta]$
0	5.2917E-014	1.6461E-003	3.1114E-002	0.0000E+000	–5.2323E-002	1.0000E+000
1	5.3642E-002	1.2743E-003	1.3259E-003	1.2687E-001	–5.2379E-002	2.5022E-001
2	1.7295E-001	4.8820E-004	1.9495E-003	2.2082E-001	–5.2452E-002	7.6656E-001
3	3.5754E-002	9.2257E-005	5.2707E-004	8.2980E-002	–5.2465E-002	1.0000E+000
4	1.7633E-002	1.0222E-005	4.8699E-004	7.6951E-002	–5.2466E-002	1.0000E+000
5	2.4675E-004	5.2545E-007	2.5285E-005	4.1250E-003	–5.2466E-002	1.0000E+000
6	1.7907E-007	4.8201E-010	2.4438E-006	3.7574E-004	–5.2466E-002	1.0000E+000
7	2.1335E-014	8.8428E-015	2.2453E-009	3.3599E-007	–5.2466E-002	1.0000E+000

Table 2 Relative errors during the mesh optimization process

Iteration	η_u^{rel} [%]	η_X^{rel} [%]	$\eta_{\mathcal{J}}^{\text{rel}}$ [%]
0	1.2475e+001	4.3465e+000	2.7233e-001
1	8.5868e+000	3.2306e+000	1.6568e-001
2	2.7892e+000	1.2628e+000	2.6839e-002
3	1.4723e+000	7.0776e-001	2.5854e-003
4	7.6473e-002	3.8580e-002	2.6867e-005
5	7.3401e-003	3.3895e-003	1.1397e-007
6	6.7417e-006	3.0294e-006	1.0000e-013
7	0	0	0

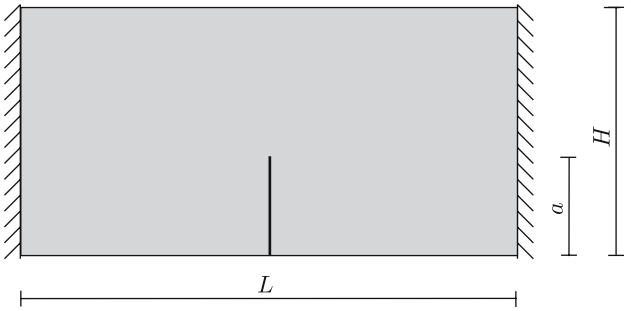


Fig. 2 Cracked plate under self-weight loading

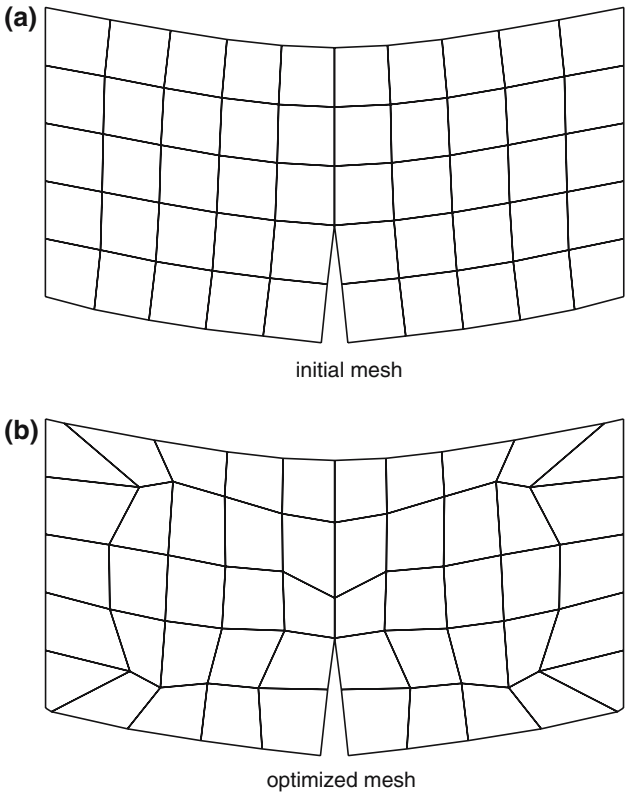


Fig. 3 Deformed configurations of the cracked plate

where $||\cdot||_{L_2}$ denotes the usual L_2-norm. The behavior of the errors during the iterations are shown in Table 2. The reduction of the error in the energy is merely marginal, but the reduction of the error in the displacement is about 12.5 %.

5.5 Example 2: cracked plate

As second example, we consider a cracked plate under self-weight loading $b_y = -5$, see Fig. 2. The dimensions of the rectangular domain are $L = 4$ and $H = 2$ as well as the crack length is $a = 0.8$. We use the same material properties like in the former example from Sect. 5.4.

All inner nodes as well as all nodal coordinates in tangential direction on the Neumann boundary Γ_{N} are chosen as design variables. Only the nodal coordinates on the Dirichlet boundary are fixed. We use the staggered solution algorithm Eq. (104) with a regularization based on the SVD and the mesh distortion control parameter (112) with $\text{tol}_\xi = 0.2$.

The solution was attained within 12 iterations. The energy decreases from -0.49541 (initial mesh) to -0.51302 (optimized mesh). The norm of the material residual $||\boldsymbol{G}||_{L_2}$ decreases from 0.0289 to 0.0104, i.e. a reduction of 64 % with respect to the initial mesh. The material residual does not vanish completely as a result of the mesh distortion control. A remeshing strategy with a change of the nodal connectivities or h-adaptivity techniques for patches with distorted elements should be considered. Furthermore, other mesh distortion control criteria, e.g. energy based, could be used in order to improve the algorithm.

As a quantity of interest, we control the vertical displacement at the crack tip $u_{y,c}$. We use a reference solution $u^*_{y,c} = -0.046105$ obtained from a fine mesh with 10,392 nodes in order to quantify the error in the displacement. The vertical displacement of the initial mesh on the crack tip is $u^0_{y,c} = -0.043173$. The relative error with respect to $u^*_{y,c}$ is given by 6.4 % For the optimized mesh follows the displacement $u_{y,c} = -0.044631$ and hence a relative error of 3.2 %. Finally, we obtain a reduction of 50 % in the relative error for the displacement at the crack tip.

6 Application to shape optimization

6.1 Problem formulation

As a second application, we consider a shape optimization problem. In order to highlight the relations between shape optimization and configurational mechanics we choose the strain energy as objective functional. This is a measure for the mean compliance of the structure. The minimization of the compliance C is equivalent to the maximization of the stiffness $S = -C$.

We consider the model problem given in Appendix A. Let

$$C = \int_{\Omega_R} W_R \, \mathrm{d}\Omega = \frac{1}{2} a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{u}) \tag{116}$$

be the strain energy of the structure and $F(\boldsymbol{s}; \boldsymbol{u})$ a linear functional associated with the external energy. The overall energy of this problem can be written in the form

$$\begin{aligned} E &= \int_{\Omega_R} W_R \, \mathrm{d}\Omega - F(\boldsymbol{s}; \boldsymbol{u}) = \frac{1}{2} a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{u}) - F(\boldsymbol{s}; \boldsymbol{u}) \\ &= -\int_{\Omega_R} W_R \, \mathrm{d}\Omega = -\frac{1}{2} a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{u}) = -C \end{aligned} \tag{117}$$

where the equilibrium equation $a(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) = F(\boldsymbol{s}; \boldsymbol{\eta}) \;\; \forall \, \boldsymbol{\eta} \in \mathcal{V}$ has been used. We are looking for the optimal design for which the structure attains a minimum of mean elastic compliance among the structures of constant volume V_0 or material cost. We use a staggered solution algorithm and we formulate the optimization problem only in terms of the design. This ends in the following problem.

Problem 7 Find $\boldsymbol{s} \in \mathcal{D}$ of the objective functional $\mathcal{J} : \mathcal{D} \to \mathbb{R}$ such that the strain energy

$$\mathcal{J}(\boldsymbol{s}) = \mathcal{J}(\boldsymbol{u}(\boldsymbol{s}), \boldsymbol{s}) = C \to \min_{\boldsymbol{s} \in \mathcal{D}} \tag{118}$$

subject to the constraints

$$\mathcal{R}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}) = 0 \tag{119}$$

$$V - V_0 = 0. \tag{120}$$

From the above definitions follow, that the variation of the objective functional, i.e. the variation of the internal energy with respect to $\boldsymbol{s}$ is the negative material residual $\mathcal{G}$, which is defined as the variation of the overall energy E with respect to $\boldsymbol{s}$, i.e.

$$E'_s = \mathcal{G}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}) \tag{121}$$

$$C'_s = -E'_s = -\mathcal{G}(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}). \tag{122}$$

Hence, the gradients which are used in the shape optimization process are the negative configurational forces.

6.2 Example 3: shape optimization of a cantilever

We consider as a simple example from shape optimization the well-known cantilever beam problem. The cantilever is clamped on the left side and loaded by a point load $f_0 = 5$ at $(L, 0)$, see Fig. 4. The initial design consists of the rectangular domain with dimensions of $L = 4$ and $H = 2$, respectively. The cantilever is modeled with the plane strain condition with $E = 1,000$ and $\nu = 0.3$. A regular finite element mesh with overall 16×8 elements is chosen to guarantee visibility of the meshes in all pictures. Standard Q4 displacement elements are utilized to ease the usage of the stated results by other researchers. We refer to [9] for the sensitivity analysis of an advanced finite element formulation.

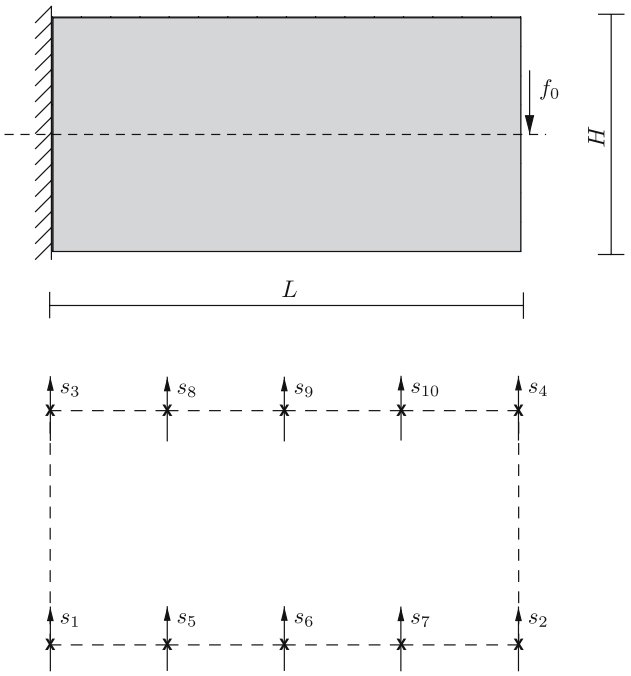


Fig. 4 System of the cantilever and geometry model with 10 design variables s_i in vertical direction. The lower and upper boundaries are modeled with Bézier curves with three internal control points (s_5–s_{10})

The optimization task is to generate the most efficient material distribution with respect to the overall stiffness of the structure. The nonlinear programming problem consists of the objective function (internal elastic energy), the constraint function (constant volume) and the geometrical design variables, i.e. we solve Problem 7. Here, the lower and upper boundaries of the structure are modeled using Bézier curves with three internal control points which are equally distributed over the length of the structure. Including the corner points, overall 10 vertical coordinates are design variables s_i, see Fig. 4. The design variables s_i are bounded by upper s_o and lower s_u side constraints

$$\begin{aligned} -10 = s_u \le s_i \le s_o = -0.25 \quad & \forall \, i = \{1, 2, 5, 6, 7\} \\ 0.25 = s_u \le s_i \le s_o = 10 \quad & \forall \, i = \{3, 4, 8, 9, 10\}. \end{aligned}$$

The optimum was attained within 19 iterations. The initial and final solutions as well as some selected iterations are shown in Fig. 6. The optimal shape leads to a uniform distribution of the material residual on the design boundary. The norm of the material residual $\boldsymbol{G}(\boldsymbol{s})$ on the design variables decreases from 0.4508 to 0.2175, i.e. a reduction of 51.8 %. The internal energy decreases from 0.4375 to 0.2869, i.e. a reduction of 34.4 %.

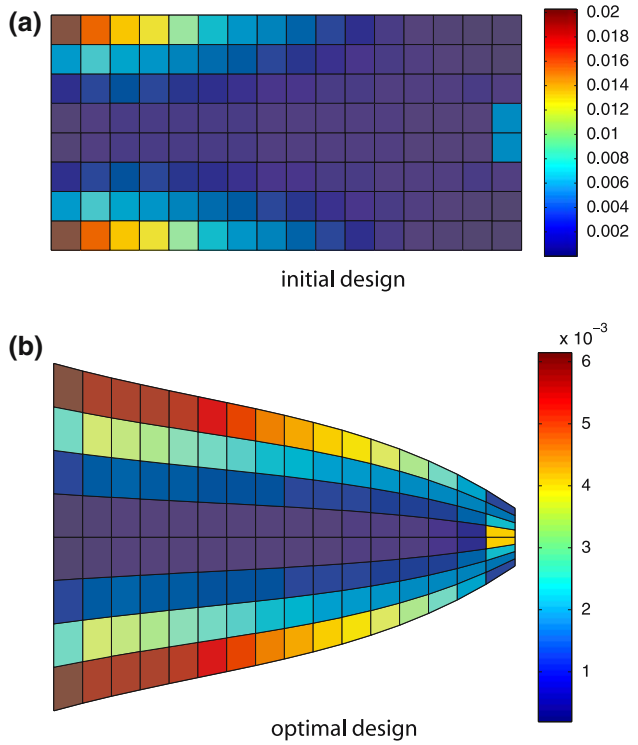


Fig. 5 Distribution of the internal energy

The distribution of the internal energy for the initial and optimized design is shown in Fig. 5. The energy distribution of the optimized shape is more smooth in comparison with the initial shape.

Remark 21 It is interesting to note, that for the compliance or internal energy optimization problem, the material residual $\mathcal{G}$ on the design boundary is the indicator in which direction the boundary has to move in order to minimize the internal energy. The optimal shape is obtained, if the material residual is uniformly distributed over the design boundary, see Fig. 6d. The material residual on the boundary does not vanish, because there is still an ambition to find a state with lower internal energy, i.e. a more stiffer structure. The side constraints for the design variables and the volume constraint avoid this movement.

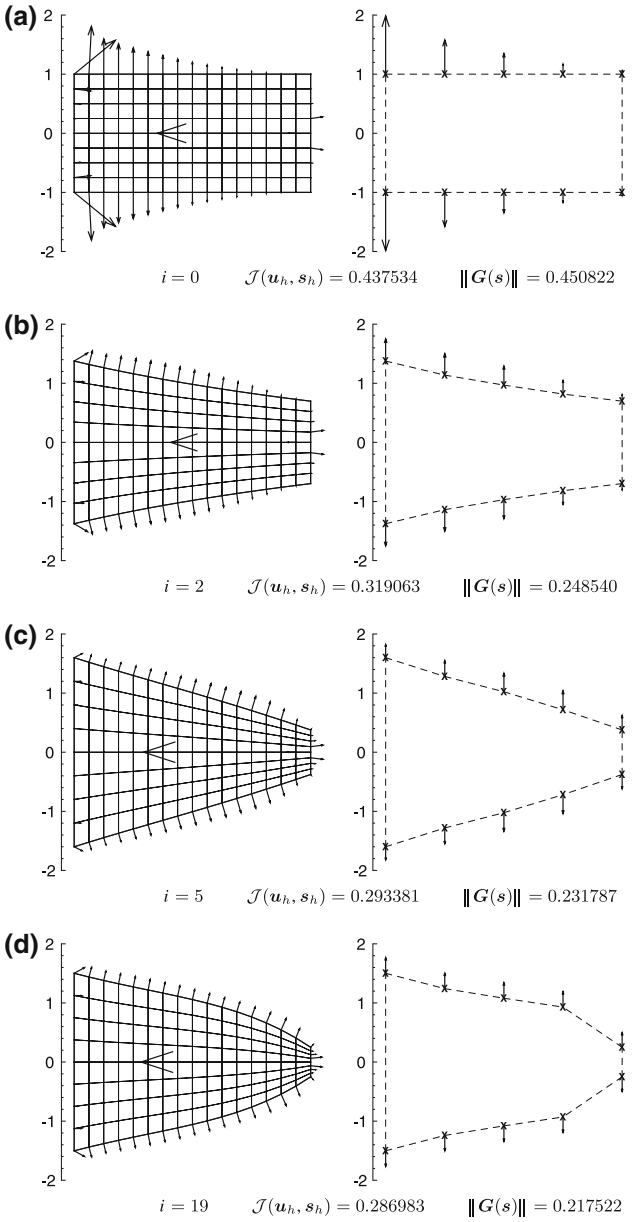


Fig. 6 Distribution of the material residuals on the mesh nodes $\boldsymbol{G}(\boldsymbol{X})$ (left side) and design variables $\boldsymbol{G}(\boldsymbol{s})$ (control points of the Bézier curves) (right side) during the optimization progress, (a) initial shape (d) final shape

6.3 Design velocity fields

The *design velocity field* in the language of structural optimization corresponds to the "displacement" in the material space. An initial domain $\Omega_R(s)$ is perturbed to a domain Ω_s by using a (time-like) design parametrization s, which parameterizes in an abstract sense the material body in the reference configuration, see Sect. 2.2. This change in the shape can be expressed by the mapping $\boldsymbol{T} : \boldsymbol{X} \to \boldsymbol{X}_s(\boldsymbol{X})$, $\boldsymbol{X} \in \Omega_R$, where $\boldsymbol{X}_s$ is the material point in the perturbed domain. By thinking of s as a time-like design parametrization, a design velocity field $\boldsymbol{V}$ can be defined as

$$\boldsymbol{V}(\boldsymbol{X}_s, s) = \frac{d\boldsymbol{X}_s}{ds} = \frac{d\boldsymbol{T}(\boldsymbol{X}, s)}{ds} = \frac{\partial \boldsymbol{T}(\boldsymbol{X}, s)}{\partial s}. \tag{123}$$

If we assume regularity in the neighborhood of the initial design $s = 0$, we have around the initial mapping point $\boldsymbol{T}(\boldsymbol{X}, 0)$ the Taylor expansion

$$\boldsymbol{T}(\boldsymbol{X}, s) = \boldsymbol{T}(\boldsymbol{X}, 0) + s \frac{\partial \boldsymbol{T}(\boldsymbol{X}, 0)}{\partial s} + \mathcal{O}. \tag{124}$$

Finally, with $\boldsymbol{T}(\boldsymbol{X}, 0) = \boldsymbol{X}$ and by ignoring high-order terms $\mathcal{O}$, we obtain the relation

$$\boldsymbol{X}_s = \boldsymbol{X} + s\,\boldsymbol{V}(\boldsymbol{X}). \tag{125}$$

The design velocity field characterizes the direction of domain variation. For a given $\boldsymbol{V}$, the change of the domain is uniquely controlled by the scalar parameter s. For more details, see e.g. [17,18] and the references therein.

For a geometry model representation of the domain, this field is generated by a variation of the underlying geometry mappings, i.e. the Bézier representation of patches describing the geometry of the structure, respectively. In the continuum domain, the velocity fields associate the geometry parametrization to the movement of material points. In the discrete case, this is the association between the geometry parametrization and the movement of mesh nodes.

We consider a vector of design variables $\boldsymbol{s} \in \mathbb{R}^m$, which parameterize the shape of the current domain Ω_R. The design velocity field corresponding to the design variable s_i is given by

$$\boldsymbol{V}_i(\boldsymbol{X}) = \frac{\partial \boldsymbol{X}}{\partial s_i}\,. \tag{126}$$

In the context of the finite element method, the design velocity field $\boldsymbol{V}_i(\boldsymbol{X})$ characterize the changes of the finite element nodal point coordinates $\boldsymbol{X}$ with respect to the changes of arbitrary design parameter s_i. For example, a variation of the vertical position of the design variables s_2 and s_5 of the lower boundary of the cantilever generate the fields shown in Fig. 7a and b, respectively. Hence, these fields reflect the influence domains of the design variations.

Furthermore, the velocity fields are also important and fundamental in the context of mesh updating and smoothing. Let $\boldsymbol{X}$ be the vector of nodal coordinates, than the new shape is obtained from

$$\boldsymbol{X}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s}) + \varepsilon \Delta \boldsymbol{X}(\boldsymbol{s}) \tag{127}$$

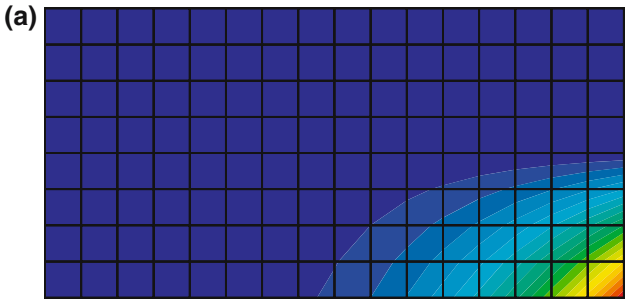
with

$$\Delta \boldsymbol{X}(\boldsymbol{s}) = \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{s}} \Delta \boldsymbol{s} = \sum_{i=1}^{m} \frac{\partial \boldsymbol{X}}{\partial s_i} \Delta s_i = \sum_{i=1}^{m} \boldsymbol{V}_i\, \Delta s_i. \tag{128}$$

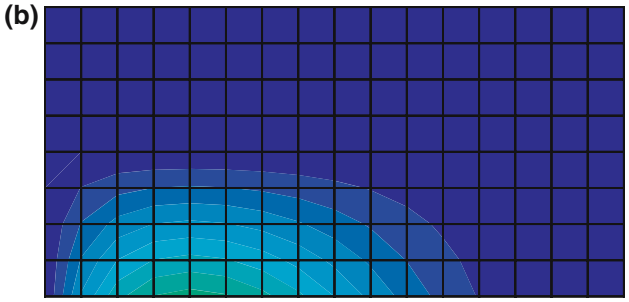
Here, ε is a step size parameter, which controls the decrease in the objective and the mesh distortion as well as $\Delta \boldsymbol{s}$ is the increment of the design variables obtained from the solution at the current iteration. This step size parameter was also used for the mesh optimization problem in Sect. 5.

The design velocity field can be multiplied with the vector field of energy variation, i.e. the material residual or configurational forces $\boldsymbol{G}$, to build up the sensitivity of the objective (the energy) with respect to a variation of the design variable s_i, i.e.

$$E'_{s_i} = \boldsymbol{G} \cdot \delta \boldsymbol{X}_i = \boldsymbol{G} \cdot \frac{\partial \boldsymbol{X}}{\partial s_i}\, \delta s_i = \boldsymbol{G} \cdot \boldsymbol{V}_i\, \delta s_i. \tag{129}$$



design variable s_2



design variable s_5

Fig. 7 Design velocity fields in vertical direction for the initial design

Hence, the design velocity fields connect the variation in the geometrical design variable s_i with the variation in the objective functional.

7 Conclusions

We have presented relations between variational design sensitivity analysis and configurational mechanics. In both disciplines, we are interested in variations of the material body and we deal with the same quantities but merely different designations. We naturally obtain the relations from configurational mechanics by using techniques from variational design sensitivity analysis applied to the energy functional of the problem. Furthermore, we have obtained sensitivity relations for the physical and material problem and with these a solution algorithm for the coupled problem. Both problems are coupled by the pseudo load operator, which play an important role for the solution of classical structural optimization problems.

With a look from the sensitivity analysis and structural optimization point of view, we have shown that the operators derived from the variations of the physical and material residuals provide much more information as only the gradients for the solution algorithms. The operators could be used for sensitivity analysis of selected quantities. Moreover, tech-

niques from structural optimization could be used for the numerical solution.

As a possible application, we have considered a mesh optimization problem. Furthermore, an example from shape optimization was investigated in order to highlight the relations between shape optimization and configurational mechanics.

Future work should address the application of the proposed framework to large systems and real structural optimization problems. More complex elastic and inelastic material laws should be considered. Furthermore, for the mesh optimization problem, the material residual is an error indicator for a non-optimal discretization. The development of practical error estimators for the discretization error as well as for the error in the design based on variational sensitivity analysis is desirable.

Appendix A: Model problem

A.1 Energy functional

We consider as a simple model problem the boundary value problem of linear elasticity. The corresponding energy functional for this problem reads

$$E(\boldsymbol{u},\boldsymbol{s}) = \int_{\Omega_R} U_R \,\mathrm{d}\Omega = \int_{\Omega_R} [W_R + V_R]\,\mathrm{d}\Omega \tag{130}$$

with

$$W_R := \frac{1}{2}\,\boldsymbol{\sigma} : \nabla^s \boldsymbol{u}, \quad V_R := -\boldsymbol{b}_R \cdot \boldsymbol{u} \tag{131}$$

where $\boldsymbol{\sigma} := \mathbb{C} : \nabla^s \boldsymbol{u} = \mathbb{C} : \boldsymbol{\varepsilon}(\boldsymbol{u})$ is the linear Cauchy stress tensor and $\boldsymbol{b}_R$ are the body forces in the reference configuration Ω_R.

A.2 Sensitivity of the energy functional

The first variation of the energy functional with respect to changes in the deformation leads to the physical residual (29)

$$\mathcal{R} = a(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta}) - F(\boldsymbol{s};\boldsymbol{\eta}) = 0 \quad \forall\, \boldsymbol{\eta} \in \mathcal{V}.$$

In linear elasticity, $a(\cdot,\cdot)$ is a bilinear form and $F(\cdot)$ is a linear functional, which are defined as

$$a(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta}) = \int_{\Omega_R} \boldsymbol{\sigma}(\boldsymbol{u}) : \nabla\boldsymbol{\eta}\,\mathrm{d}\Omega \tag{132}$$

$$F(\boldsymbol{s};\boldsymbol{\eta}) = \int_{\Omega_R} \boldsymbol{b}_R \cdot \boldsymbol{\eta}\,\mathrm{d}\Omega. \tag{133}$$

The corresponding strong form is an elliptic partial differential equation for the state and reads

$$\boldsymbol{A}(\boldsymbol{u},\boldsymbol{s}) = \operatorname{div}\boldsymbol{\sigma} + \boldsymbol{b}_R = \boldsymbol{0} \quad \text{in } \Omega_R. \tag{134}$$

In the same manner, the first variation of the energy functional with respect to changes in the design leads to the material residual (30),

$$\mathcal{G} = b(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi}) - L(\boldsymbol{s};\boldsymbol{\psi}) = 0 \quad \forall\, \boldsymbol{\psi} \in \mathcal{D}$$

which contains the semi-linear form $b(\cdot,\cdot)$ and the linear functional $L(\cdot)$. For our model problem these quantities are given by

$$b(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi}) = \int_{\Omega_R} (\boldsymbol{\Sigma} + V_R\boldsymbol{1}) : \nabla\boldsymbol{\psi}\,\mathrm{d}\Omega \tag{135}$$

$$L(\boldsymbol{s};\boldsymbol{\psi}) = \int_{\Omega_R} \boldsymbol{b}_R^{inh} \cdot \boldsymbol{\psi}\,\mathrm{d}\Omega \tag{136}$$

where

$$\boldsymbol{\Sigma} := W_R\boldsymbol{1} - (\nabla\boldsymbol{u})^{\mathrm{T}}\boldsymbol{\sigma} \tag{137}$$

is the energy momentum or Eshelby tensor, here in terms of linear elasticity. Furthermore, in the case that the strain energy $W(\boldsymbol{X};\nabla^s\boldsymbol{u})$ depends explicitly on the position $\boldsymbol{X}$, the inhomogeneity force $\boldsymbol{b}_R^{inh}$ is given by the explicit derivative of W, i.e.

$$\boldsymbol{b}_R^{inh} := -\left.\frac{\partial W_R}{\partial \boldsymbol{X}}\right|_{\text{expl.}}. \tag{138}$$

The corresponding strong form of the configurational force equilibrium or pseudomomentum equation reads

$$\boldsymbol{B}(\boldsymbol{u},\boldsymbol{s}) = \operatorname{div}\boldsymbol{\Sigma} + \boldsymbol{g}_R = \boldsymbol{0} \quad \text{in } \Omega_R \tag{139}$$

where

$$\boldsymbol{g}_R = -(\nabla\boldsymbol{u})^{\mathrm{T}}\boldsymbol{b}_R + \boldsymbol{b}_R^{inh} \tag{140}$$

are the configurational volume forces and $\boldsymbol{B}(\boldsymbol{u},\boldsymbol{s})$ is the differential operator for the material problem.

Remark 22 For a homogeneous elastic body without any material inhomogeneity the second term of $\boldsymbol{g}$ vanishes, because in this case the strain energy W depends not explicitly on the position $\boldsymbol{X}$.

A.3 Sensitivity of the physical residual

The variation of the physical residual $\mathcal{R}$ with respect to $\boldsymbol{u}$ leads to the physical stiffness (42)

$$k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) = a'_u(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u}) = \int_{\Omega_R} \nabla\boldsymbol{\eta} : \mathbb{C} : \operatorname{sym}\{\nabla\delta\boldsymbol{u}\}\,\mathrm{d}\Omega \tag{141}$$

whereas we have assumed, that the linear functional $F(\cdot)$ is independent of the deformation, i.e. $F'_u = 0$.

The contributions of the pseudo load operator (43) are given by

$$\begin{aligned} a'_s(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}, \delta\boldsymbol{s}) = &- \int_{\Omega_R} \nabla\boldsymbol{\eta} : \mathbb{C} : \operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\delta\boldsymbol{s}\}\, \mathrm{d}\Omega \\ &- \int_{\Omega_R} \boldsymbol{\sigma} : \nabla\boldsymbol{\eta}\, \nabla\delta\boldsymbol{s}\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} \boldsymbol{\sigma} : \nabla\boldsymbol{\eta} \operatorname{div}\delta\boldsymbol{s}\, \mathrm{d}\Omega \end{aligned} \tag{142}$$

and

$$F'_s(\boldsymbol{s}; \boldsymbol{\eta}, \delta\boldsymbol{s}) = \int_{\Omega_R} \boldsymbol{b}_R \cdot \boldsymbol{\eta} \operatorname{div}\delta\boldsymbol{s}\, \mathrm{d}\Omega \tag{143}$$

whereas we have assumed, that $\boldsymbol{b}_R$ is independent of the design (e.g. constant). Hence, the pseudo-load operator (43) becomes

$$\begin{aligned} p(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}, \delta\boldsymbol{s}) = &- \int_{\Omega_R} \nabla\boldsymbol{\eta} : \mathbb{C} : \operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\delta\boldsymbol{s}\}\, \mathrm{d}\Omega \\ &- \int_{\Omega_R} \boldsymbol{\sigma} : \nabla\boldsymbol{\eta}\, \nabla\delta\boldsymbol{s}\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} [\boldsymbol{\sigma} : \nabla\boldsymbol{\eta} - \boldsymbol{b}_R \cdot \boldsymbol{\eta}]\, \mathbf{1} : \nabla\delta\boldsymbol{s}\, \mathrm{d}\Omega, \end{aligned} \tag{144}$$

where $\operatorname{div}\delta\boldsymbol{s} = \mathbf{1} : \nabla\delta\boldsymbol{s}$ has been used.

In order to obtain more compact and demonstrative forms of the above expressions we consider the linear strain tensor

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^{\mathrm{T}}) = \operatorname{sym}\{\nabla\boldsymbol{u}\}. \tag{145}$$

The variations of $\boldsymbol{\varepsilon}$ with respect to $\boldsymbol{u}$ and $\boldsymbol{s}$ as well as the mixed variations are given by

$$\boldsymbol{\varepsilon}'_u(\boldsymbol{\eta}) = \frac{1}{2}(\nabla\boldsymbol{\eta} + \nabla\boldsymbol{\eta}^{\mathrm{T}}) = \operatorname{sym}\{\nabla\boldsymbol{\eta}\} = \boldsymbol{\varepsilon}(\boldsymbol{\eta}) \tag{146}$$

$$\begin{aligned} \boldsymbol{\varepsilon}'_s(\boldsymbol{u}, \boldsymbol{\psi}) &= -\frac{1}{2}(\nabla\boldsymbol{u}\, \nabla\boldsymbol{\psi} + (\nabla\boldsymbol{u}\, \nabla\boldsymbol{\psi})^{\mathrm{T}}) \\ &= -\operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\boldsymbol{\psi}\} \end{aligned} \tag{147}$$

$$\begin{aligned} \boldsymbol{\varepsilon}''_{us}(\boldsymbol{\eta}, \boldsymbol{\psi}) &= -\frac{1}{2}(\nabla\boldsymbol{\eta}\, \nabla\boldsymbol{\psi} + (\nabla\boldsymbol{\eta}\, \nabla\boldsymbol{\psi})^{\mathrm{T}}) \\ &= -\operatorname{sym}\{\nabla\boldsymbol{\eta}\, \nabla\boldsymbol{\psi}\} \end{aligned} \tag{148}$$

$$\boldsymbol{\varepsilon}''_{su}(\boldsymbol{\psi}, \boldsymbol{\eta}) = \boldsymbol{\varepsilon}''_{us}(\boldsymbol{\eta}, \boldsymbol{\psi}) = -\operatorname{sym}\{\nabla\boldsymbol{\eta}\, \nabla\boldsymbol{\psi}\} \tag{149}$$

The stiffness operator and the pseudo load operator could now be expressed as

$$\begin{aligned} k(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}, \delta\boldsymbol{u}) &= \int_{\Omega_R} \boldsymbol{\varepsilon}'_u(\boldsymbol{\eta}) : \mathbb{C} : \boldsymbol{\varepsilon}'_u(\delta\boldsymbol{u})\, \mathrm{d}\Omega \\ &= \int_{\Omega_R} \boldsymbol{\varepsilon}(\boldsymbol{\eta}) : \mathbb{C} : \boldsymbol{\varepsilon}(\delta\boldsymbol{u})\, \mathrm{d}\Omega \end{aligned} \tag{150}$$

$$\begin{aligned} p(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\eta}, \delta\boldsymbol{s}) = &\int_{\Omega_R} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}''_{us}(\boldsymbol{\eta}, \delta\boldsymbol{s})\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} \boldsymbol{\varepsilon}'_u(\boldsymbol{\eta}) : \mathbb{C} : \boldsymbol{\varepsilon}'_s(\boldsymbol{u}, \delta\boldsymbol{s})\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} [\boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{\eta}) - \boldsymbol{b}_R \cdot \boldsymbol{\eta}]\, \mathbf{1} : \nabla\delta\boldsymbol{s}\, \mathrm{d}\Omega. \end{aligned} \tag{151}$$

A.4 Sensitivity of the material residual

As a result of the symmetry of the pseudo load operator for the physical (43) and material problem (51), i.e.

$$p(\boldsymbol{u}, \boldsymbol{s}; \cdot, \delta\boldsymbol{u}) = p(\boldsymbol{u}, \boldsymbol{s}; \delta\boldsymbol{u}, \cdot) \tag{152}$$

we have to calculate only the variation of $\mathcal{G}$ with respect to $\boldsymbol{s}$. If we consider a homogeneous elastic body without any material inhomogeneity, the variation of $L(\cdot)$ vanishes identically, i.e. $L'_s = 0$. Therefore, only the variation $b'_s(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}, \delta\boldsymbol{s})$ contribute to the material tangent operator $d(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}, \delta\boldsymbol{s})$. A straightforward calculation leads to

$$\begin{aligned} &d(\boldsymbol{u}, \boldsymbol{s}; \boldsymbol{\psi}, \delta\boldsymbol{s}) \\ &= \int_{\Omega_R} \boldsymbol{\sigma} : \operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\delta\boldsymbol{s}\, \nabla\boldsymbol{\psi}\}\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} \boldsymbol{\sigma} : \operatorname{sym}\{\nabla\boldsymbol{u}\nabla\boldsymbol{\psi}\, \nabla\delta\boldsymbol{s}\}\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} \operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\boldsymbol{\psi}\} : \mathbb{C} : \operatorname{sym}\{\nabla\boldsymbol{u}\, \nabla\delta\boldsymbol{s}\}\, \mathrm{d}\Omega \\ &- \int_{\Omega_R} \boldsymbol{\sigma} : \operatorname{sym}\{\nabla\boldsymbol{u}\nabla\boldsymbol{\psi}\} \operatorname{div}\delta\boldsymbol{s}\, \mathrm{d}\Omega \\ &- \int_{\Omega_R} \boldsymbol{\sigma} : \operatorname{sym}\{\nabla\boldsymbol{u}\nabla\delta\boldsymbol{s}\} \operatorname{div}\boldsymbol{\psi}\, \mathrm{d}\Omega \\ &+ \int_{\Omega_R} W_R \left[\operatorname{div}\boldsymbol{\psi} \operatorname{div}\delta\boldsymbol{s} - \mathbf{1} : \nabla\,\psi\nabla\delta\boldsymbol{s}\right] \mathrm{d}\Omega \end{aligned} \tag{153}$$

The second variation of the linear strain tensor $\boldsymbol{\varepsilon}$ with respect to $\boldsymbol{s}$ is given by

$$\begin{aligned}\boldsymbol{\varepsilon}''_{ss}(\boldsymbol{u},\boldsymbol{\psi},\boldsymbol{\chi}) &= \frac{1}{2}\Big(\nabla\boldsymbol{u}\,\nabla\boldsymbol{\chi}\,\nabla\boldsymbol{\psi}+\nabla\boldsymbol{u}\,\nabla\boldsymbol{\psi}\,\nabla\boldsymbol{\chi}\\ &\quad+(\nabla\boldsymbol{u}\,\nabla\boldsymbol{\chi}\,\nabla\boldsymbol{\psi})^{\mathrm{T}}+(\nabla\boldsymbol{u}\,\nabla\boldsymbol{\psi}\,\nabla\boldsymbol{\chi})^{\mathrm{T}}\Big)\\ &= \operatorname{sym}\{\nabla\boldsymbol{u}\,\nabla\boldsymbol{\chi}\,\nabla\boldsymbol{\psi}+\nabla\boldsymbol{u}\,\nabla\boldsymbol{\psi}\,\nabla\boldsymbol{\chi}\}. \qquad (154)\end{aligned}$$

Finally, from this follow directly

$$\begin{aligned}&d(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{s})\\ &=\int_{\Omega_R}\boldsymbol{\sigma}:\boldsymbol{\varepsilon}''_{ss}(\boldsymbol{u},\boldsymbol{\psi},\delta\boldsymbol{s})\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}\boldsymbol{\varepsilon}'_{s}(\boldsymbol{u},\boldsymbol{\psi}):\mathbb{C}:\boldsymbol{\varepsilon}'_{s}(\boldsymbol{u},\delta\boldsymbol{s})\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}\boldsymbol{\sigma}:\boldsymbol{\varepsilon}'_{s}(\boldsymbol{u},\boldsymbol{\psi})\operatorname{div}\delta\boldsymbol{s}\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}\boldsymbol{\sigma}:\boldsymbol{\varepsilon}'_{s}(\boldsymbol{u},\delta\boldsymbol{s})\operatorname{div}\boldsymbol{\psi}\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}W_R\left[\operatorname{div}\boldsymbol{\psi}\operatorname{div}\delta\boldsymbol{s}-\mathbf{1}:\nabla\psi\nabla\delta\boldsymbol{s}\right]\mathrm{d}\Omega \qquad (155)\end{aligned}$$

$$\begin{aligned}&p(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\psi},\delta\boldsymbol{u})\\ &=\int_{\Omega_R}\boldsymbol{\sigma}:\boldsymbol{\varepsilon}''_{su}(\boldsymbol{\psi},\delta\boldsymbol{u})\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}\boldsymbol{\varepsilon}'_{s}(\boldsymbol{u},\boldsymbol{\psi}):\mathbb{C}:\boldsymbol{\varepsilon}'_{u}(\delta\boldsymbol{u})\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}[\,\boldsymbol{\sigma}:\boldsymbol{\varepsilon}(\delta\boldsymbol{u})-\boldsymbol{b}_R\cdot\delta\boldsymbol{u}]\,\mathbf{1}:\nabla\boldsymbol{\psi}\,\mathrm{d}\Omega. \qquad (156)\end{aligned}$$

Remark 23 Due to the linearity of the chosen model problem, the stiffness operator $k(\cdot;\cdot)$ in Eq. (150) contains only one part without a dependency on $\boldsymbol{u}$. In the general geometrical nonlinear case with the Green–Lagrange strain tensor $\boldsymbol{E}$ and the second Piola–Kirchhoff stress tensor $\boldsymbol{S}$ is the tangent stiffness operator given by

$$\begin{aligned}&k(\boldsymbol{u},\boldsymbol{s};\boldsymbol{\eta},\delta\boldsymbol{u})\\ &=\int_{\Omega_R}\boldsymbol{S}:\boldsymbol{E}''_{uu}(\boldsymbol{\eta},\boldsymbol{\delta u})\,\mathrm{d}\Omega\\ &\quad+\int_{\Omega_R}\boldsymbol{E}'_{u}(\boldsymbol{u},\boldsymbol{\eta}):\mathbb{C}:\boldsymbol{E}'_{u}(\boldsymbol{u},\boldsymbol{\delta u})\,\mathrm{d}\Omega. \qquad (157)\end{aligned}$$

The first term is often called *initial stress* term and the second term contains the incremental constitutive tensor $\mathbb{C}$ as well as different variations of the Green-Lagrange strain tensor [53]. It is interesting to note that the first and the second term of the operators $p(\cdot;\cdot)$ and $d(\cdot;\cdot)$ have the same structure like the usual tangent stiffness operator of the physical problem for the geometrical nonlinear case (157) with different variations of the strain tensor $\boldsymbol{\varepsilon}$. Additionally, we have further terms in these operators as a result of variations of the domain Ω_R. An explizit formulation of all operators for nonlinear hyperelasticity is given in [36].

References

1. Askes H, Kuhl E, Steinmann P (2004) An ALE formulation based on spatial and material settings of continuum mechanics. Part 2: Classification and applications. Comput Methods Appl Mech Eng 193:4223–4245
2. Askes H, Bargmann S, Kuhl E, Steinmann P (2005) Structural optimisation by simultaneous equilibration of spatial and material forces. Commun Numer Meth Eng 21:433–442
3. Barthold FJ (2002) Zur Kontinuumsmechanik inverser Geometrieprobleme. Habilitationsschrift, Braunschweiger Schriften zur Mechanik 44-2002, TU Braunschweig, URL http://hdl.handle.net/2003/23095
4. Barthold FJ (2003) A structural optimization viewpoint on configurational mechanics. Proc Appl Math Mech 3:246–247
5. Barthold FJ (2005) On structural optimisation and configurational mechanics. In: Steinmann P, Maugin GA (eds) Mechanics of material forces, vol 11. Springer, Heidelberg, pp 219–228
6. Barthold FJ (2007) Remarks on variational shape sensitivity analysis based on local coordinates. Eng Anal Bound Elem (in press)
7. Barthold FJ, Mesecke S (1999) Remarks on computing the energy release rate and its sensitivities. In: Mota Soares C, Mota Soares C, Freitas M (eds) Mechanics of composite materials and structures, nato science series, Serie E: Applied Science, vol 361. Kluwer, Dordrecht, pp 341–350
8. Barthold FJ, Stein E (1996) A continuum mechanical based formulation of the variational sensitivity analysis in structural optimization. Part I. Anal Struct Optim 11(1/2):29–42
9. Barthold FJ, Wiechmann K (2006) A comparison of displacement and mixed finite element formulations for variational design sensitivity analysis. In: Mota Soares CA et al (ed) III European conference on computational mechanics, Paper 2003
10. Bendsøe MP, Sigmund O (2003) Topology optimization-theory, methods and applications. Springer, Heidelberg
11. Benzi M, Golub GH (2004) A preconditioner for generalized saddle point problems. Siam J Matrix Anal Appl 26(1):20–41
12. Bertram A (1989) Axiomatische Einführung in die Kontinuumsmechanik. BI-Wissenschaftsverlag, Mannheim
13. Braun M (1997) Configurational forces induced by finite-element discretisation. Proc Estonian Acad Sci Phys Math 46:24–31
14. Braun M (2005) Structural optimization by material forces. In: Steinmann P, Maugin GA (eds) Mechanics of material forces, vol 11. Springer, Heidelberg, pp 211–218
15. Carpenter WC, Zendegui S (1982) Optimum nodal locations for a finite element idealization. Eng Optim 5:215–221
16. Carroll WE, Barker RM (1973) A theorem for optimum finite-element idealizations. Int J Solids Struct 9:883–895
17. Choi KK, Kim NH (2005a) Structural sensitivity analysis and optimization 1—Linear systems. Mechanical Engineering Series. Springer, Berlin
18. Choi KK, Kim NH (2005b) Structural sensitivity analysis and optimization 2—Nonlinear systems and applications, Mechanical Engineering Series. Springer, Berlin

19. Dems K, Mróz Z (1978) Multiparameter structural shape optimization by the finite element method. Int J Numer Methods Eng 13(2):247–263
20. Engl HW, Hanke M, Neubauer A (2000) Regularization of inverse problems. Kluver, Dordrecht
21. Eshelby JD (1951) The force on an elastic singularity. Philos Trans R Soc Lond 244:87–112
22. Eshelby JD (1975) The elastic energy-momentum tensor. J Elast 5:321–335
23. Golub GH, Loan CFV (1996) Matrix computations, 3rd edn. The Johns Hopkins University Press, Baltimore
24. Govindjee S, Mihalic PA (1996) Computational methods for inverse finite elastostatics. Comput Methods Appl Mech Eng 136(1–2):47–57
25. Govindjee S, Mihalic PA (1998) Computational methods for inverse deformations in quasi-incompressible finite elasticity. Int J Numer Meth Eng 43(5):821–838
26. Gurtin ME (2000) Configurational forces as basic concepts of continuum physics. Springer, New York
27. Kalpakides VK, Balassas KG (2005) The inverse deformation mapping in the finite element method. Philos Mag 85(33–35):4257–4275
28. Kalpakides VK, Maugin G (eds) (2004) Configurational mechanics. A.A. Balkema, Amsterdam
29. Kamat MP (1993) Structural optimization: status and promise. American Institute of Aeronautics and Astronautics, Washington
30. Kienzler R, Herrmann G (2000) Mechanics in material space. Springer, Berlin
31. Kuhl E, Askes H, Steinmann P (2004) An ALE formulation based on spatial and material settings of continuum mechanics. Part 1: Generic hyperelastic formulation. Comput Methods Appl Mech Eng 193:4207–4222
32. Marsden JE, Hughes TJR (1994) Mathematical foundations of elasticity. Dover, New York
33. Materna D, Barthold FJ (2006a) Coherence of structural optimization and configurational mechanics. Proc Appl Math Mech 6(1):245–246. doi:10.1002/pamm.200610103
34. Materna D, Barthold FJ (2006b) Relations between structural optimization and configurational mechanics with applications to mesh optimization. In: Sienz J, Querin OM, Toropov VV, Gosling PD (eds) Proceedings of the 6th ASMO-UK/ISSMO Conference on Engineering Design Optimization, St Edmund Hall, Oxford, UK, University of Leeds, UK, pp 173–181
35. Materna D, Barthold FJ (2007a) The use of singular value decomposition in sensitivity analysis and mesh optimization. (submitted)
36. Materna D, Barthold FJ (2007b) Variational design sensitivity analysis in the context of structural optimization and configurational mechanics. Int J Fract (in press)
37. Maugin GA (1993) Material inhomogeneities in elasticity. Chapman & Hall, London
38. Maugin GA (1995) Material forces: concepts and applications. Appl Mech Rev 48:213–245
39. Maugin GA, Trimarco C (1992) Pseudomomentum and material forces in nonlinear elasticity: variational formulations and application to brittle fracture. Acta Mech 94:1–28
40. McNeice GM, Marcal PE (1973) Optimization of finite element grids based on minimum potential energy. ASME J Eng Ind 95(1):186–190
41. Mosler J, Ortiz M (2006) On the numerical modeling of variational arbitrary lagrangian-eulerian (vale) formulations. Int J Numer Meth Eng 67:1272–1289
42. Mueller R, Maugin GA (2002) On material forces and finite element discretizations. Comput Mech 29:52–60
43. Mueller R, Kolling S, Gross D (2002) On configurational forces in the context of the finite element method. Int J Numer Meth Eng 53:1557–1574
44. Mueller R, Gross D, Maugin GA (2004) Use of material forces in adaptive finite element methods. Comput Mech 33:421–434
45. Nocedal J, Wright SJ (1999) Numerical optimization. Springer, Heidelberg
46. Noll W (1972) A new mathematical theory of simple materials. Arch Ration Mech Anal 48(1):1–50
47. Steinmann P (2000) Application of material forces to hyperelastic fracture mechanics. I. Continuum mechanical settings. Int J Solids Struct 37:7371–7391
48. Steinmann P, Maugin GA (eds) (2005) Mechanics of material forces. Springer, Heidelberg
49. Sussman T, Bathe KJ (1985) The gradient of the finite element variational indicator with respect to nodal point coordinates: an explicit calculation and applications in fracture mechanics and mesh optimization. Int J Numer Meth Eng 21(4):763–774
50. Thoutireddy P (2003) Variational arbitrary Lagrangian-Eulerian method. Phd thesis, California Institute of Technology, Passadena, CA
51. Thoutireddy P, Ortiz M (2004) A variational r-adaption and shape-optimization method for finite-deformation elasticity. Int J Numer Meth Eng 61(1):1–21
52. Truesdell C, Noll W (2004) The nonlinear field theories of mechanics, 3rd edn. Springer, Berlin
53. Wriggers P (2001) Nichtlineare finite-element-methoden. Springer, Heidelberg
54. Zhang S, Belegundu A (1993) Mesh distortion control in shape optimization. AIAA J 31(7):1360–1362