Shape Optimization of Biomorphic Ceramics with Microstructures by Homogenization Modeling

Ronald H.W. Hoppe[1,2] and Svetozara I. Petrova[1,3]

[1] Institute of Mathematics, University of Augsburg, 86159 Augsburg.
hoppe@math.uni-augsburg.de, petrova@math.uni-augsburg.de
[2] Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA
[3] Institute for Parallel Processing, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

Summary. We consider the modeling, simulation, and optimization of microstructural cellular biomorphic ceramics obtained by biotemplating. This is a process in biomimetics, a recently emerged discipline in materials science where engineers try to mimick or use biological materials for the design of innovative technological devices and systems. In particular, we focus on the shape optimization of microcellular silicon carbide ceramic materials derived from naturally grown wood. The mechanical behavior of the final ceramics is largely determined by the geometry of its microstructure which can be very precisely tuned during the biotemplating process. Our ultimate goal is to determine these microstructural details in such a way that an optimal mechanical performance is achieved with respect to merit criteria depending on the specific application. Within the shape optimization problem the state variables are the displacements subject to the underlying elasticity equations, and the design variables are the geometrical quantities determining the microstructure. Since a resolution of the microstructure is numerically cost-prohibitive, we use the homogenization approach, assuming periodically distributed microcells. Adaptive mesh-refinement techniques based on reliable and efficient a posteriori error estimators are applied in the microstructure to compute the homogenized elasticity coefficients. The shape optimization problem on the macroscopic homogenized model is solved by primal-dual Newton-type interior-point methods. Various numerical experiments are presented and discussed.

1 Introduction

Biomimetics, also called bionics or biomimicry, is a discipline in materials science that has recently attracted a lot of attention. It allows the cost-effective production of high performance technological devices and systems by either mimicking or using naturally grown biological structures (cf., e.g.,

[Eli00, HDL02, GA88]). In contrast to engineering materials, biological structures exhibit a hierarchically built anatomy, developed and optimized in a long-term genetic process. Their inherent cellular, open porous morphology can be used for liquid or gaseous infiltration and subsequently for high temperature reaction processes. A specific example of such a naturally grown biological material is wood which exhibits an anisotropic, porous morphology with excellent strength at low density, high stiffness, elasticity, and damage tolerance. Typical feature of the wood structure is the system of the tracheidal cells which provide the transportation path for water and minerals within the living tree. This open porous system is accessible for infiltration of various metals.

A recent idea in biomimetical applications is to take advantage of naturally grown wood in the production of high performance ceramics to be used as filters and catalysts in chemical processing, heat insulation structures, thermally and mechanically loaded lightweight structures, and medical implants (for instance, for bone substitution). In particular, silicon carbide (SiC) is known as a material that is not only suitable in microelectronical applications due to its bandwidth structure but also useful in mechanical and high temperature applications with regard to its excellent thermomechanical properties. Since wood essentially consists of carbon (C), the idea is to use it as a basic material for the production of highly porous ceramics. Among the large variety of ceramic composites, new biomorphic cellular silicon carbide ceramics from wood were recently produced and investigated, see [GLK98a, GLK98b, OT*95, VSG02]. The conversion of naturally grown wood to highly porous SiC ceramics is done by a process called *biotemplating* which includes two processing steps, illustrated in Fig. 1.1.

Biological porous carbonized preforms (also called $C-$templates) can be derived from different wood structures by drying and high-temperature pyrolysis at temperatures between 800 and 1800^oC and used as templates for infiltrations by gaseous or liquid silicon (Si) to form SiC and SiSiC-ceramics, respectively. During high-temperature processing, the microstructural properties of the bioorganic preforms were retained, so that a one-to-one reproduc-

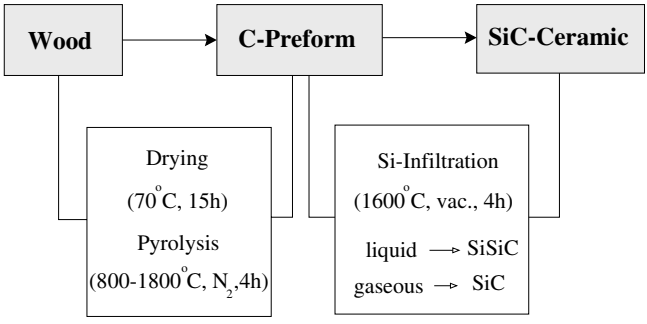

Fig. 1.1. Processing scheme of SiC-ceramics from wood

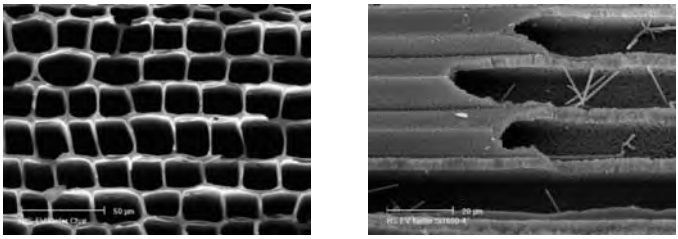

Fig. 1.2. SiC-ceramic derived from pine wood a) radial direction; b) axial direction

tion of the original wood structure was obtained, see Fig. 1.2. The resulting cellular composites show low density, high specific strength, and excellent high temperature stability.

The geometry of the final ceramics, i.e., the widths and lengths of the different layers forming the struts, can be determined very precisely by an appropriate tuning of the process parameters. This raises the question how to choose these microstructural geometrical data in order to achieve an optimal performance with respect to a prespecified merit criterion depending on the specific application. From a mathematical point of view, this issue represents a shape optimization problem where the *state variables* are subject to the underlying elasticity equations and the microstructural data serve as *design variables*. As far as the solution of such a shape optimization problem is concerned, the resolution of the microstructure is cost-prohibitive with respect to both computational time and storage. Therefore, the idea is to perform homogenization, assuming a periodically distributed microstructure, and to apply the optimization to the homogenized model.

In this study, we focus both on the homogenization process and on the application of state-of-the-art optimization strategies for the numerical solution of the shape optimization problem under consideration. The remaining of the paper is organized as follows: In Sect. 2, we describe in detail the homogenization technique that provides a macromechanical model where the components of the homogenized elasticity tensor reflect the microstructural details. Section 3 deals with the setting of our shape optimization problem. In Sect. 4, we present a primal-dual Newton interior-point method and in Sect. 5 we comment on the numerical solution of the condensed primal-dual system. Section 6 concerns adaptive grid-refinement procedures based on a posteriori error estimators. In particular, we use the *Zienkiewicz-Zhu* error estimator proposed in [ZZ87]. Iterative solution techniques for the homogenized elasticity equation and the microcell problem are discussed in Sect. 7. Various numerical results are given in Sect. 8.

2 Homogenized computational model

In this section, we briefly explain the derivation of the homogenized computational model on the macroscale by using the asymptotic homogenization theory. Homogenization has been successfully used in the last three decades for solving multi-scale problems on computational regions occupied by heterogeneous microstructural materials (see, cf., [BP84, BLP78, JKO94, SP80]).

Let $\Omega \subset \mathcal{R}^d,\ d = 2, 3$, be a domain occupied by a heterogeneous material with microstructures of periodically distributed constituents. Suppose that the boundary of Ω, denoted by Γ, consists of a prescribed displacement boundary Γ_D (meas $\Gamma_D > 0$) and a prescribed traction boundary Γ_T, such that $\Gamma = \Gamma_D \cup \Gamma_T,\ \Gamma_D \cap \Gamma_T = \emptyset$. Let $\boldsymbol{b}$ be the body force, $\bar{\boldsymbol{u}}$ be the prescribed displacement on Γ_D, and $\bar{\boldsymbol{t}}$ be the prescribed traction on Γ_T.

The homogenized model for our original heterogeneous material occupying the domain Ω, $\Omega \subset \mathcal{R}^d,\ d = 2, 3$, is illustrated in Fig. 2.1. The main idea of the homogenization is to replace the heterogeneous material by an equivalent homogenized material, extracting the information for the material properties of the various microstructural constituents (or different phases).

The microscopic and macroscopic models are considered simultaneously supposing a strong scale separation, i.e., a large gap in length scale between the macroscopic component and the microstructure. In practical applications the microscopic length scales are orders of magnitude smaller than the physical macroscopic length scale. A main assumption in the homogenization approach is that the original heterogeneous material workpiece is composed of periodically distributed microstructures of various constituents. To couple properly the micro- and macro-scales, we choose a representative volume element (RVE) or a unit microstructure.

Consider a stationary microstructure with a geometrically simple tracheidal periodicity cell $Y = [0, 1]^d,\ d = 2, 3$, (see Fig. 2.2) consisting of an outer layer of carbon (C), interior layer of silicon carbide (SiC), and a

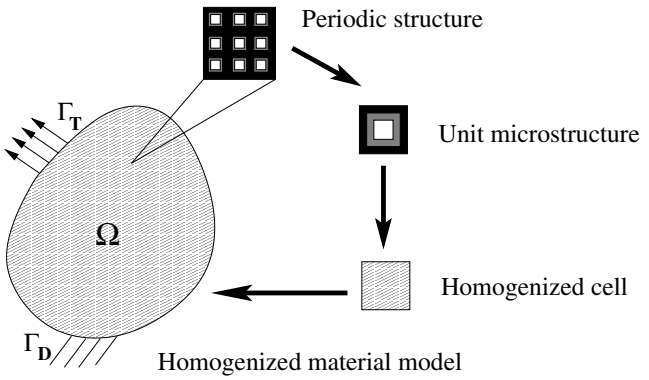

Fig. 2.1. The macroscopic homogenized material model

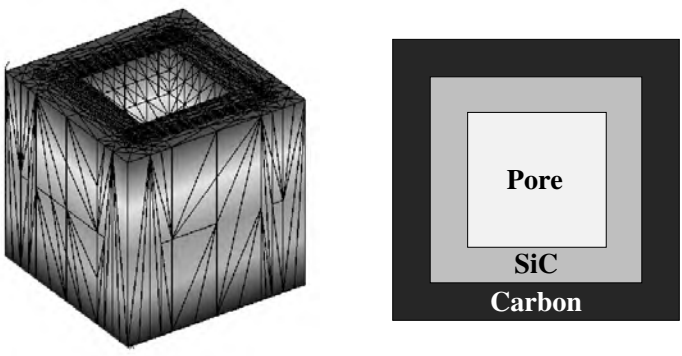

Fig. 2.2. a) 3-D unit periodicity cell Y, b) 2-D unit periodicity cell $Y = P \cup SiC \cup C$

centered pore channel (P, no material). We introduce two space variables $\boldsymbol{x}$ (macroscopic/slow variable) and $\boldsymbol{y}$ (microscopic/fast variable) and denote by ε, $\boldsymbol{y} = \boldsymbol{x}/\varepsilon$, $\varepsilon \ll 1$, the scale parameter (dimensionless number) which, in fact, represents the periodicity under the assumption that ε is very small with respect to the size of Ω, i.e., there exists a large scale gap between the microstructure and the macroscopic component.

The parameter ε allows us to define macrofunctions in terms of the microstructural behavior and vice versa. Thus, for any state function $f(\boldsymbol{y}) := f(\boldsymbol{x}/\varepsilon)$, one can compute the spatial derivatives by using the following differentiation rule

$$\frac{d}{d\boldsymbol{x}} f\left(\boldsymbol{x}, \frac{\boldsymbol{x}}{\varepsilon}\right) = \frac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial \boldsymbol{x}} + \varepsilon^{-1} \frac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial \boldsymbol{y}}.$$

Consider the following elasticity equation defined in the microstructure Y

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{F} \qquad \text{in } \ Y \tag{2.1}$$

with a load vector $\mathbf{F}$. Here, $\boldsymbol{\sigma}$ is the microscopic symmetric stress, $\boldsymbol{u} \in \boldsymbol{H}^1(Y)$, is the corresponding displacement at point $\boldsymbol{y} \in Y$, and $\boldsymbol{e}$ is the microscopic symmetric strain with components

$$e_{ij}(\boldsymbol{u}(\boldsymbol{y})) = \frac{1}{2}\left(\frac{\partial u_i(\boldsymbol{y})}{\partial y_j} + \frac{\partial u_j(\boldsymbol{y})}{\partial y_i}\right). \tag{2.2}$$

The problem (2.1) is subject to periodic boundary conditions on the outer part of ∂Y, Neumann boundary conditions around the pore, and continuity conditions $[\boldsymbol{u}] = 0$ and $[\boldsymbol{\sigma} \cdot \boldsymbol{n}] = 0$ on the interfaces between the different phases, see Fig. 2.2. The symbol $[\ ]$ denotes the jump of the function across the corresponding interface with a normal vector $\boldsymbol{n}$ (cf., e.g., [BP84]).

Assuming linearly elastic constituents, the unit microstructure is governed by the Hooke law $\boldsymbol{\sigma} = \boldsymbol{E} : \boldsymbol{e}$ with componentwise $(i, j, k, l = 1, \ldots, d)$ constitutive relations as follows

$$\sigma_{ij}(\boldsymbol{u}) = E_{ijkl}(\boldsymbol{y})\, e_{kl}(\boldsymbol{u}(\boldsymbol{y})). \tag{2.3}$$

Here, we adopt the Einstein convention of a summation on repeated indices. The 4-th order elasticity (also called plain-stress) tensor $\boldsymbol{E}(\boldsymbol{y})$ with components $E_{ijkl}(\boldsymbol{y})$ characterizes the behavior of the material at point $\boldsymbol{y}$ and depends on material constants like Young's modulus and Poisson's ratio. Note that $\boldsymbol{E}(\boldsymbol{y})$ is zero if $\boldsymbol{y}$ is located in the porous subdomain of the microstructure and coincides with the elasticity tensor of the material if $\boldsymbol{y}$ is located in the corresponding microstructural constituent. The elasticity tensor is symmetric in the following sense

$$E_{ijkl} = E_{jikl} = E_{ijlk} = E_{klij} \tag{2.4}$$

and satisfies the following ellipticity conditions

$$E_{ijkl}\,\chi_{ij}\chi_{kl} \geq c\,\chi_{ij}^2, \qquad \forall\, \chi_{ij} = \chi_{ji},$$

for a constant $c > 0$ (cf., e.g., [BP84, BLP78, JKO94]).

Denote by $\boldsymbol{u}_\varepsilon(\boldsymbol{x}) := \boldsymbol{u}(\boldsymbol{x}/\varepsilon)$ the unknown macroscopic displacement vector and consider the following family of elasticity problems

$$-\nabla \cdot \boldsymbol{\sigma}_\varepsilon \,=\, \boldsymbol{b} \qquad \text{in } \Omega, \tag{2.5}$$

subject to a macroscopic body force $\boldsymbol{b}$ and a macroscopic surface traction $\boldsymbol{t}$ applied to the portion $\Gamma_T \subset \partial\Omega$. Here, $\boldsymbol{\sigma}_\varepsilon(\boldsymbol{u}_\varepsilon) := \boldsymbol{E}_\varepsilon(\boldsymbol{x})\boldsymbol{e}(\boldsymbol{u}_\varepsilon(\boldsymbol{x}))$ is the stress tensor for $\boldsymbol{x} \in \Omega$ and $\boldsymbol{E}_\varepsilon(\boldsymbol{x}) := \boldsymbol{E}(\boldsymbol{x}/\varepsilon) = \boldsymbol{E}(\boldsymbol{y})$ is the piecewise constant elasticity tensor defined in Y. Following, for instance, [BLP78] for the basic concepts of the homogenization method, the unknown displacement vector is expanded asymptotically as

$$\boldsymbol{u}_\varepsilon(\boldsymbol{x}) = \boldsymbol{u}^{(0)}(\boldsymbol{x},\boldsymbol{y}) + \varepsilon\,\boldsymbol{u}^{(1)}(\boldsymbol{x},\boldsymbol{y}) + \varepsilon^2\,\boldsymbol{u}^{(2)}(\boldsymbol{x},\boldsymbol{y}) + \dots\,, \qquad \boldsymbol{y} = \boldsymbol{x}/\varepsilon, \tag{2.6}$$

where $\boldsymbol{u}^{(i)}(\boldsymbol{x},\boldsymbol{y})$, $i \geq 0$, are $Y-$periodic in $\boldsymbol{y}$, i.e., take equal values on opposite sides of Y. Consider the space $H := \{\boldsymbol{u} | \boldsymbol{u} \in \boldsymbol{H}^1(\Omega),\ \boldsymbol{u} = 0 \text{ on } \Gamma_D\}$. Under the assumptions of symmetry and ellipticity of the elasticity coefficients, it was shown in the homogenization theory that the sequence $\{\boldsymbol{u}_\varepsilon\}$ of solutions of (2.5) tends weakly in H as $\varepsilon \to 0$ to a function $\boldsymbol{u}^{(0)}(\boldsymbol{x}) \in H$, the solution of the following macroscopic homogenized problem with a constant elasticity tensor.

$$-\nabla \cdot \boldsymbol{\sigma} = \boldsymbol{b} \qquad \text{in } \Omega, \tag{2.7}$$

where $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{u}^{(0)}) := \boldsymbol{E}^H \boldsymbol{e}(\boldsymbol{u}^{(0)}(\boldsymbol{x}))$, $\boldsymbol{x} \in \Omega$, and $\boldsymbol{E}^H$ stands for the homogenized elasticity tensor. Note that $\boldsymbol{u}^{(0)}(\boldsymbol{x})$ depends only on the macroscopic variable $\boldsymbol{x}$ and is independent of the microscopic scale $\boldsymbol{y}$. The leading term $\boldsymbol{u}^{(0)}$ in (2.6) is called a macroscopic displacement and the remaining terms $\boldsymbol{u}^{(i)}$, $i > 0$, are considered as perturbed displacements.

The homogenization method requires to find periodic functions $\boldsymbol{\xi}^{kl}$ satisfying the following problem in a weak formulation to be solved in the microscopic unit cell

$$\int_Y E_{ijpq}(\boldsymbol{y})\frac{\partial \xi_p^{kl}}{\partial y_q}\frac{\partial \phi_i}{\partial y_j}\,dY = \int_Y E_{ijkl}(\boldsymbol{y})\frac{\partial \phi_i}{\partial y_j}\,dY, \tag{2.8}$$

where $\boldsymbol{\phi} \in \boldsymbol{H}^1(Y)$ is an arbitrary $Y-$periodic variational function. The function $\boldsymbol{\xi}^{kl}$, also referred to as the characteristic displacement, is found by solving (2.8) in Y with periodic boundary conditions. After computing $\boldsymbol{\xi}^{kl}$, one defines the homogenized coefficients by the following formulas (we refer to [BP84, BLP78, JKO94] for details)

$$E^H_{ijkl} = \frac{1}{|Y|}\int_Y \left(E_{ijkl}(\boldsymbol{y}) - E_{ijpq}(\boldsymbol{y})\frac{\partial \xi_p^{kl}}{\partial y_q}\right) dY. \tag{2.9}$$

Due to the symmetry conditions (2.4), the 4-th order homogenized elasticity tensor $\boldsymbol{E}^H = (E^H_{ijkl})$ can be written as a symmetric and usually a dense matrix. In the case $d = 2$ it is a 3×3 matrix and has the form

$$\mathbf{E}^H = \begin{pmatrix} E^H_{1111} & E^H_{1122} & E^H_{1112} \\ & E^H_{2222} & E^H_{2212} \\ \text{SYM} & & E^H_{1212} \end{pmatrix}. \tag{2.10}$$

The 3-d homogenized tensor can be written, respectively, as a 6×6 matrix

$$\mathbf{E}^H = \begin{pmatrix} E^H_{1111} & E^H_{1122} & E^H_{1133} & E^H_{1112} & E^H_{1123} & E^H_{1113} \\ & E^H_{2222} & E^H_{2233} & E^H_{2212} & E^H_{2223} & E^H_{2213} \\ & & E^H_{3333} & E^H_{3312} & E^H_{3323} & E^H_{3313} \\ & & & E^H_{1212} & E^H_{1223} & E^H_{1213} \\ & & & & E^H_{2323} & E^H_{2313} \\ \text{SYM} & & & & & E^H_{1313} \end{pmatrix}. \tag{2.11}$$

The computation of the homogenized elasticity coefficients can be done analytically for some specific geometries as, for instance, layered materials or checkerboard structures. In case of more complicated microstructures, the computation of E^H_{ijkl} has to be done numerically through a suitable microscopic modeling.

Once the constant homogenized coefficients from (2.9) are computed, one comes up with the homogenized macroscopic equation (2.7) given in a weak form as follows

$$\int_\Omega E^H_{ijkl}\frac{\partial u^0_k}{\partial x_l}\frac{\partial v_i}{\partial x_j}d\Omega = \int_\Omega \boldsymbol{b}\cdot\boldsymbol{v}\,d\Omega + \int_{\Gamma_T}\bar{\boldsymbol{t}}\cdot\boldsymbol{v}\,d\Gamma, \qquad \forall\ \boldsymbol{v}\in H, \tag{2.12}$$

where $\boldsymbol{u}^{(0)}(\boldsymbol{x}) := \boldsymbol{u}^{(0)}(\boldsymbol{x},\boldsymbol{y})$ is the homogenized solution occurring in (2.6).

3 Shape optimization by primal-dual methods

Structural optimization has recently become of increasing interest in computer aided design and optimization of composite structures in materials science (cf.,

e.g., [Ben95] and the references therein). A typical problem of structural optimization is to minimize a function (called *objective*, *cost* or *criterion* function) over a set of geometrical or behavioral requirements (called *constraints*). The set of structural parameters includes the so-called *state* and *design* parameters, and the problem consists in computing optimal values of the design parameters, such that they minimize the specific objective function. Sizing, shape, and topology optimization problems are different types in structural optimization. Detailed classification of these problems is given, for instance, in [OT83]. In the sizing problems, the goal is to find the optimal thickness distribution of a given material structure. The main difficulty in shape optimization problems arises from the fact that the geometry of a structure is a design variable which means, in particular, that the discretization model associated with the structure has to be changed in the process of optimization, see [All02, SZ92]. In the topology optimization of solid structures we are interested in the determination of the optimal placement of material in space, i.e., one has to determine which points of space are material and which points should remain void (no material). Hence, the main goal of these problems is to find the location of holes and the connectivity of the domain, see [BS03].

Our goal is to optimize mechanical performances of the ceramic composites described in Sect. 1 (such as the compliance or the bending strength) taking into account technological and problem specific constraints on the state and design parameters. Denote the state variables $\boldsymbol{u} = (u_1, ..., u_N)^T$, which are the nodal values of the components of the discrete displacement vector, and the design variables $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_M)^T$ chosen as the microstructural data determining the geometry of the periodicity cell (widths and lengths of the different materials layers forming the cell walls, see Fig. 2.2). Since the geometrical properties of the final ceramics are not fixed but can be changed and precisely tuned within the processing, we focus on shape optimization of our microcellular SiC ceramic materials. Depending on the specific application, the objective functional $J = J(\boldsymbol{u}, \boldsymbol{\alpha})$ can be chosen according to the following criteria:

- mechanical properties (minimum compliance),
- loading (bending, tension, torsion),
- thermal properties (shock resistance),
- technological properties (minimum weight),
- economical properties (cheapness).

For simplicity, we consider the mean compliance of the structure defined as

$$J(\boldsymbol{u}, \boldsymbol{\alpha}) = \int_{\Omega} \boldsymbol{b} \cdot \boldsymbol{u} \, d\Omega + \int_{\Gamma_T} \bar{\boldsymbol{t}} \cdot \boldsymbol{u} \, d\Gamma, \tag{3.1}$$

Our shape optimization problem reads: Find $(\boldsymbol{u}, \boldsymbol{\alpha}) \in \mathcal{R}^N \times \mathcal{R}^M$ such that

$$J(\boldsymbol{u}, \boldsymbol{\alpha}) = \inf_{\boldsymbol{v}, \boldsymbol{\beta}} J(\boldsymbol{v}, \boldsymbol{\beta}), \tag{3.2}$$

subjected to the following equality and inequality constraints

$$A(\boldsymbol{\alpha})\boldsymbol{u} = \boldsymbol{f}, \qquad g(\boldsymbol{\alpha}) := \sum_{i=1}^{M} \alpha_i = C, \qquad \alpha_{\min}\,\bar{\boldsymbol{e}} \le \boldsymbol{\alpha} \le \alpha_{\max}\,\bar{\boldsymbol{e}}, \tag{3.3}$$

where $J(\boldsymbol{u}, \boldsymbol{\alpha})$ is defined by (3.1), $A(\boldsymbol{\alpha})$ is the stiffness matrix corresponding to the homogenized equilibrium equation (2.12), $\boldsymbol{u}$ is the discrete homogenized displacement vector, $\boldsymbol{f}$ is the discrete load vector, and $\bar{\boldsymbol{e}} = (1, 1, \ldots, 1) \in \mathcal{R}^M$. Note that $\alpha_{\min}$ and $\alpha_{\max}$ are technologically motivated lower and upper bounds for the design parameters. In the case of unit microstructure Y, we take the limits $\alpha_{\min} = 0$, $\alpha_{\max} = 0.5$, and $0 < C \le 0.5$.

4 Primal–dual Newton interior–point method

In the optimization algorithm, we are typically faced with constrained nonconvex nonlinear minimization problems with both equality and inequality constraints on the state variables and design parameters. For the discretized optimization problem we use the primal-dual Newton interior-point methods, recently a topic of intensive research [BHN99, ET*86, FGW02, GOW98, HPS02, HP04b, VS99]. The main idea of these methods is to generate iteratively approximations of the solution which strictly satisfy the inequality constraints. Details are given in this section.

4.1 General nonlinear optimization problem

We consider the following general constrained nonlinear nonconvex programming problem with both equality and inequality constraints

$$\min_{\boldsymbol{x} \in \mathcal{R}^n} f(\boldsymbol{x}), \tag{4.1}$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = 0, \qquad \boldsymbol{g}(\boldsymbol{x}) \ge 0, \tag{4.2}$$

where $f : \mathcal{R}^n \to \mathcal{R}$, $\boldsymbol{h} : \mathcal{R}^n \to \mathcal{R}^m, m < n$, and $\boldsymbol{g} : \mathcal{R}^n \to \mathcal{R}^l$ are assumed to be twice Lipschitz continuously differentiable. Note that the constraints (4.2) have to be understood componentwise.

The *Lagrangian function* associated with (4.1)–(4.2) is defined by

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = f(\boldsymbol{x}) + \boldsymbol{y}^T \boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{z}^T \boldsymbol{g}(\boldsymbol{x}), \tag{4.3}$$

where $\boldsymbol{y} \in \mathcal{R}^m$ and $\boldsymbol{z} \in \mathcal{R}^l$ are the Lagrange multipliers for the equality and inequality constraints, respectively.

The first-order Karush-Kuhn-Tucker (KKT) necessary conditions for optimality of (4.1)–(4.2) read

$$\nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})=0,\ \boldsymbol{h}(\boldsymbol{x})=0,\ \boldsymbol{g}(\boldsymbol{x})\geq 0,\ Z\,\boldsymbol{g}(\boldsymbol{x})=0,\ \boldsymbol{z}\geq 0,$$

where

$$\nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})=\nabla f(\boldsymbol{x})+\sum_{i=1}^{m}y_i\nabla h_i(\boldsymbol{x})-\sum_{i=1}^{l}z_i\nabla g_i(\boldsymbol{x}) \tag{4.4}$$

is the gradient of the Lagrangian function and Z is the diagonal matrix with a diagonal $\boldsymbol{z}$. We also consider the Hessian of the Lagrangian with respect to $\boldsymbol{x}$ defined by

$$\nabla_x^2\mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})=\nabla^2 f(\boldsymbol{x})+\sum_{i=1}^{m}y_i\nabla^2 h_i(\boldsymbol{x})-\sum_{i=1}^{l}z_i\nabla^2 g_i(\boldsymbol{x}), \tag{4.5}$$

where $\nabla^2 f(\boldsymbol{x}),\ \nabla^2 h_i(\boldsymbol{x}),\ 1\leq i\leq m,\ \nabla^2 g_i(\boldsymbol{x}), 1\leq i\leq l$ stand for the Hessians of $f(\boldsymbol{x}), h_i(\boldsymbol{x})$, and $g_i(x)$, respectively. Denote by

$$\mathcal{A}(\boldsymbol{x})=\{i,\ g_i(\boldsymbol{x})=0,\ i=1,\ldots,l\}$$

the set of all indices for which the inequality constraints are equal to zero at $\boldsymbol{x}$. We are interested in finding local minimizers of our optimization problem (4.1)–(4.2). Assume that at least one such point $\boldsymbol{x}^*$ exists satisfying the conditions:

- **Feasibility.** $\boldsymbol{h}(\boldsymbol{x}^*)=0$ and $\boldsymbol{g}(\boldsymbol{x}^*)\geq 0$.
- **Regularity.** The set $\{\nabla h_1(\boldsymbol{x}^*),\ldots,\nabla h_m(\boldsymbol{x}^*)\}\cup\{\nabla g_i(\boldsymbol{x}^*), i\in\mathcal{A}(\boldsymbol{x}^*)\}$ of gradients of equality and active inequality constraints is linearly independent.
- **Smoothness.** The Hessian matrices $\nabla^2 f(\boldsymbol{x}),\ \nabla^2 h_i(\boldsymbol{x}),\ 1\leq i\leq m$, and $\nabla^2 g_i(\boldsymbol{x}),\ 1\leq i\leq l$, exist and are locally Lipschitz continuous at $\boldsymbol{x}^*$.
- **Second-order sufficiency condition.** $\boldsymbol{\eta}^T\nabla^2_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}^*)\boldsymbol{\eta}>0$ for all vectors $\boldsymbol{\eta}\neq 0$ satisfying $\nabla h_i(\boldsymbol{x}^*)^T\boldsymbol{\eta}=0,\ 1\leq i\leq m$, and $\nabla g_i(\boldsymbol{x}^*)^T\eta=0,\ i\in\mathcal{A}(\boldsymbol{x}^*)$.
- **Strict complementarity.** $z_i^*>0$ if $g_i(\boldsymbol{x}^*)=0,\ 1\leq i\leq l$.

Well-known approaches from the optimization theory for handling problems with inequality constraints are, for instance, the slack variable approach, the active set strategy, and the logarithmic barrier function approach. Each of these approaches results in a nonlinear programming problem with only equality constraints. For example, in the first approach, the constraint $\boldsymbol{g}(\boldsymbol{x})\geq 0$ can be replaced by $\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{s}=0, \boldsymbol{s}\geq 0$ by adding a nonnegative slack variable to each of the inequality constraints. Transformation of the original inequality problem into an equality one, by adding slacks, have been a frequently applied tool in scientific computations in the past twenty years and recently used in cf., [BHN99, ET*86, VS99]. The introduction of slack variables is associated with a small amount of additional work and storage, since they do not enter

the objective function and are constrained by simple bounds. The second, active set strategy, approach in nonlinear programming is directly related to the idea of the simplex method in linear programming. At each iterative step from this approach, applying, for example, Newton's method, one has to define which constraints are active at the solution and treat them as equality constraints by ignoring the others. The third approach was used in our practical implementations and we explain it in detail in the next subsection.

4.2 Logarithmic barrier interior-point method

The logarithmic barrier function method was first introduced in [Fri55] and later on popularized by [FM68] in the late sixties of the last century. The basic idea of this method is to replace the optimization problem (4.1)–(4.2) with the following equality constrained optimization problem

$$\min_{\boldsymbol{x}\in\mathcal{R}^n} \beta^{(\rho)}(\boldsymbol{x}), \tag{4.6}$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = 0, \tag{4.7}$$

where ρ is a positive scalar, called *barrier parameter*, and

$$\beta^{(\rho)}(\boldsymbol{x}) = f(\boldsymbol{x}) - \rho \sum_{i=1}^{l} \log\, g_i(\boldsymbol{x}) \tag{4.8}$$

is often referred to as a *barrier function*. To insure existence of the logarithmic terms in (4.8) we implicitly require $g_i(\boldsymbol{x}) > 0,\ 1 \le i \le l$. In such a way, we get a family of subproblems depending on ρ for which it is well-known that under the assumptions conditions from Sect. 4.1 the solution of (4.6)–(4.7) converges to a solution of the original problem (4.1)–(4.2) as ρ decreases to zero (cf., [FM68]). This method obviously is an interior-point method since it keeps the sequence of iterating solutions strictly feasible with respect to the inequality constraints. Note that the logarithmic terms serve as a barrier and result in finding a solution $\boldsymbol{x}^{(\rho)}$ such that $g(\boldsymbol{x}^{(\rho)}) > 0$. The solution points $\boldsymbol{x}^{(\rho)}$ parameterized by ρ define the so-called *central path* or also called *barrier trajectory*.

The gradient of (4.8) is given by

$$\nabla\beta^{(\rho)}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) - \sum_{i=1}^{l} \frac{\rho}{g_i(\boldsymbol{x})} \nabla g_i(\boldsymbol{x})$$

and the Hessian of $\beta^{(\rho)}(x)$ is defined by

$$\nabla^2\beta^{(\rho)}(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) - \sum_{i=1}^{l} \frac{\rho}{g_i(\boldsymbol{x})} \nabla^2 g_i(\boldsymbol{x}) + \sum_{i=1}^{l} \frac{\rho}{g_i^2(\boldsymbol{x})} \nabla g_i(\boldsymbol{x}) (\nabla g_i(\boldsymbol{x}))^T. \tag{4.9}$$

The Lagrangian function associated with (4.6)–(4.7) is

$$\mathcal{L}^{(\rho)}(\boldsymbol{x},\boldsymbol{y}) = \beta^{(\rho)}(\boldsymbol{x}) + \boldsymbol{y}^T\boldsymbol{h}(\boldsymbol{x}) = f(\boldsymbol{x}) - \rho\sum_{i=1}^{l}\log\, g_i(\boldsymbol{x}) + \boldsymbol{y}^T\boldsymbol{h}(\boldsymbol{x})$$

and the gradient of $\mathcal{L}^{(\rho)}(\boldsymbol{x},\boldsymbol{y})$ with respect to $\boldsymbol{x}$ is given by

$$\nabla_{\boldsymbol{x}}\mathcal{L}^{(\rho)}(\boldsymbol{x},\boldsymbol{y}) = \nabla f(\boldsymbol{x}) - \sum_{i=1}^{l}\frac{\rho}{g_i(\boldsymbol{x})}\nabla g_i(\boldsymbol{x}) + \sum_{i=1}^{m} y_i\nabla h_i(\boldsymbol{x}). \tag{4.10}$$

The logarithmic barrier function method consists now of generating a sequence of iterative solutions $\{\boldsymbol{x}\} = \{\boldsymbol{x}^{(\rho)}\}$, local minimizers of the equality constrained subproblems, with $\rho > 0$ decreasing at each iteration. Taking into account the first-order optimality conditions and especially $\nabla_{\boldsymbol{x}}\mathcal{L}^{(\rho)}(\boldsymbol{x},\boldsymbol{y}) = 0$, we see that convergence of $\{\boldsymbol{x}^{(\rho)}\}$ to an optimal solution $\boldsymbol{x}^*$ requires that

$$\lim_{\rho\to 0} y_i^{(\rho)} = y_i^*,\ 1 \le i \le m \qquad \text{and} \qquad \lim_{\rho\to 0}\frac{\rho}{g_i(\boldsymbol{x}^{(\rho)})} = z_i^*,\ 1 \le i \le l, \tag{4.11}$$

where $\{y_i^*\}$ and $\{z_i^*\}$ are the Lagrange multipliers associated with the equality and inequality constraints $g_i(\boldsymbol{x}^{(\rho)}) > 0$, respectively. From $g_i(x^{(\rho)}) \to 0$ and the second relation in (4.11) we get $\rho/g_i^2(\boldsymbol{x}^{(\rho)}) \to \infty$ and hence, the Hessian of the logarithmic barrier function (4.9) would become arbitrarily large. Comparing now relations (4.4) and (4.10), we see that $\rho/g_i(\boldsymbol{x}^{(\rho)})$ serves as a Lagrange multiplier for the inequality constraints. Thus, we can introduce an auxiliary variable $z_i = z_i^{(\rho)} = \rho/g_i(\boldsymbol{x}^{(\rho)}),\ 1 \le i \le l$ which can also be written in the form $z_i^{(\rho)} g_i(\boldsymbol{x}^{(\rho)}) = \rho$. The last relation is usually called *perturbed complementarity* and can be used as a remedy, so that the differentiation will not create ill-conditioning.

We formulate now the perturbed KKT conditions for the logarithmic barrier function problem (4.6)–(4.7), namely

$$\nabla f(\boldsymbol{x}) + \nabla\boldsymbol{h}(\boldsymbol{x})\boldsymbol{y} - \nabla\boldsymbol{g}(\boldsymbol{x})\boldsymbol{z} = 0,\ \boldsymbol{h}(\boldsymbol{x}) = 0,\ Z\boldsymbol{g}(\boldsymbol{x}) = \rho\bar{\boldsymbol{e}},\ \boldsymbol{g}(\boldsymbol{x}) > 0. \tag{4.12}$$

In matrix-vector notations, (4.12) results in the following nonlinear equation with $n + m + l$ components

$$\boldsymbol{F}^{(\rho)}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = 0 \qquad \text{with}\ \ \boldsymbol{F}^{(\rho)}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \begin{pmatrix} \boldsymbol{t} + J_{\text{eq}}^T\boldsymbol{y} - J_{\text{in}}^T\boldsymbol{z} \\ \boldsymbol{h} \\ G\boldsymbol{z} - \rho\,\bar{\boldsymbol{e}} \end{pmatrix}, \tag{4.13}$$

where $\boldsymbol{F}^{(\rho)} = \nabla\mathcal{L}^{(\rho)}$ is the gradient of the Lagrangian function with respect to $\boldsymbol{x},\boldsymbol{y}$, and $\boldsymbol{z}$; $\boldsymbol{t} = \nabla f(\boldsymbol{x})$ is the gradient of the objective function, J_{eq} is the Jacobian $m \times n$ matrix of the equality constraints $\boldsymbol{h}(\boldsymbol{x}) = 0$ and J_{in} is the Jacobian $l \times n$ matrix of the inequality constraints $\boldsymbol{g}(\boldsymbol{x}) \ge 0$. In the

last equation of (4.13) we have denoted $G = \mathrm{diag}(g_i),\ g_i > 0,\ 1 \le i \le l$ and $\bar{\boldsymbol{e}} = (1, 1, \ldots, 1)^T$. Note that at each iteration we have three sets of unknowns: the primal variable $\boldsymbol{x}$, the dual variable $\boldsymbol{y}$, and the perturbed complementarity variable $\boldsymbol{z}$ which we consider independently.

Denote the unknown solution by $\boldsymbol{\Phi} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ and apply the Newton method to the nonlinear system (4.13) to find the increments $\Delta\boldsymbol{\Phi} = (\Delta\boldsymbol{x}, \Delta\boldsymbol{y}, \Delta\boldsymbol{z})$, namely

$$K\Delta\boldsymbol{\Phi} = -\boldsymbol{F}^{(\rho)}(\boldsymbol{\Phi}), \tag{4.14}$$

which is often referred to as a *primal-dual system*. The vector $\Delta\boldsymbol{\Phi}$ is called *search direction*. The so-called *primal-dual matrix* $K = \left(\boldsymbol{F}^{(\rho)}\right)'(\boldsymbol{\Phi})$ of second derivatives of the Lagrangian function is defined as follows

$$K = \begin{pmatrix} H & J_{\mathrm{eq}}^T & -J_{\mathrm{in}}^T \\ J_{\mathrm{eq}} & 0 & 0 \\ ZJ_{\mathrm{in}} & 0 & G \end{pmatrix}, \tag{4.15}$$

where the Hessian of the Lagrangian function $H = \nabla_x^2\mathcal{L}$ is given by (4.5). Note that the matrix K is sparse, nonsymmetric, independent of ρ, and usually well-conditioned in a sense that its condition number is limited when $\rho \to 0$ (see [Wri98] for more details). In the case of convex optimization (i.e., convex objective function $f(\boldsymbol{x})$, linear equality constraints $\boldsymbol{h}(\boldsymbol{x})$, and concave inequality constraints $\boldsymbol{g}(\boldsymbol{x})$), the Hessian matrix H is positive semidefinite. The properties of the Hessian matrix for inequality constrained optimization problem with logarithmic barrier function method are discussed in [FGW02].

One possible way for solving (4.14) is to symmetrize K taking into account the fact that Z and G are diagonal matrices. This method is proposed in [FGS96] and results in the following symmetric matrix

$$\hat{K} = \begin{pmatrix} H & J_{\mathrm{eq}}^T & -J_{\mathrm{in}}^T \\ J_{\mathrm{eq}} & 0 & 0 \\ -J_{\mathrm{in}} & 0 & -Z^{-1}G \end{pmatrix},$$

which is strongly ill-conditioned with some diagonal elements becoming unbounded as $\rho \to 0$. In particular, for the active inequality constraints, the diagonal entries of $Z^{-1}G$ go to zero, and for the inactive constraints they go to infinity. As the iterates converge, the ill-conditioning of K increases, but it was shown in [FGS96] that the primal-dual solution of the optimization problem is actually independent of the size of the large diagonal elements and can be found by using, for instance, a symmetric indefinite factorization of the primal-dual system.

Another alternative way for solving (4.14) which we use in our practical applications is to eliminate the (1,3) block of (4.15), i.e., due to $\boldsymbol{g}(\boldsymbol{x}) > 0$, we eliminate $\Delta\boldsymbol{z}$ from the third equation of (4.14)

$$\Delta\boldsymbol{z} = -\boldsymbol{z} + G^{-1}(\rho\,\bar{\boldsymbol{e}} - ZJ_{\mathrm{in}}\Delta\boldsymbol{x}) \tag{4.16}$$

and replace it in the first equation. This method produces a symmetric linear system with $n + m$ equations of the form

$$\begin{pmatrix} \tilde{H} & J_{\text{eq}}^T \\ J_{\text{eq}} & 0 \end{pmatrix} \begin{pmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \end{pmatrix} = - \begin{pmatrix} \boldsymbol{t} + J_{\text{eq}}^T \boldsymbol{y} - \rho J_{\text{in}}^T G^{-1} \bar{\boldsymbol{e}} \\ \boldsymbol{h} \end{pmatrix}, \tag{4.17}$$

where $\tilde{H} = H + J_{\text{in}}^T G^{-1} Z J_{\text{in}}$ is often referred to as a *condensed* primal-dual Hessian and (4.17) is called a *condensed* primal-dual system. A detailed analysis of the properties of the condensed primal-dual matrix can be found in [Wri98] where it was shown that the inherent ill-conditioning of the reduced primal-dual matrix is usually benign and does not influence the accuracy of the solution.

Various methods for solving (4.17) and finding the primal-dual steps $(\Delta \boldsymbol{x}, \Delta \boldsymbol{y})$ are proposed in the literature (cg., e.g., [ET*86, GOW98, Wri98]). Note that one needs a reliable and efficient solver of (4.17), since the condensed primal-dual system is solved at every iteration of the optimization loop. In practice, we apply transforming iterations (see [Wit89]) to find the increments. This method will be explained in more detail in Sect. 5.

After finding the solution of (4.17), the algorithm proceeds iteratively from an initial point $(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{z}^{(0)})$ through a sequence of points determined from the search directions described by (4.16) and (4.17) as follows

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_x^{(k)} \Delta \boldsymbol{x}, \ \boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \alpha_y^{(k)} \Delta \boldsymbol{y}, \ \boldsymbol{z}^{(k+1)} = \boldsymbol{z}^{(k)} + \alpha_z^{(k)} \Delta \boldsymbol{z}.$$

The parameters $\alpha_x^{(k)}, \alpha_y^{(k)}, \alpha_z^{(k)} \in (0, 1]$ are called *steplengths* and their choice at each iteration is a critical feature of the algorithm to find a local minimizer of the optimization problem.

4.3 Merit functions. Computing the steplengths

In all optimization algorithms it is important to have a reasonable way of deciding whether the new iterate is better than the previous one, i.e., it is essential to measure appropriately the progress in finding a local solution. Merit functions of different types have been a subject of great interest over the past years (see, e.g., [ET*86, GOW98, Wri98]). The main idea of a merit function is to ensure simultaneously a progress toward a local minimizer and toward feasibility. The method of choosing $\alpha^{(k)}$ at each iteration becomes more complicated in general nonlinear programming problems as it is well-known that the Newton method may diverge when the initial estimate of the solution is bad.

Two versions of the Newton method can be applied, namely, the *trust-region* and the *line-search* approach. The first method has recently been applied in, e.g., [BHN99]. Typical for this method is to find a step $\boldsymbol{d}^{(k)}$ which is restricted to a set, called the *trust region*. This set is practically obtained

by limiting $|\boldsymbol{d}^{(k)}| \leq r^{(k)}$, where $r^{(k)}$ is the trust region radius. At each iteration, $r^{(k)}$ is updated according to how successful the step has been. For instance, if the a priori chosen merit function M decreases, we accept the step $\boldsymbol{d}^{(k)}$, update the solution $\boldsymbol{\Phi}^{(k+1)} = \boldsymbol{\Phi}^{(k)} + \boldsymbol{d}^{(k)}$ and possibly increase the trust region radius $r^{(k)}$. Otherwise, we decrease $r^{(k)}$ by a damping factor, e.g., $r^{(k)} = r^{(k)}/2$ and compute again the step $\boldsymbol{d}^{(k)}$.

We apply the second variant of the Newton method, the line-search approach. Once the solution $\Delta\boldsymbol{\Phi}^{(k)}$ of (4.14) has been determined, we find a steplength $\alpha^{(k)} > 0$ such that $\boldsymbol{\Phi}^{(k+1)} = \boldsymbol{\Phi}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{\Phi}^{(k)}$ measuring a progress in minimization at each iteration and reducing the merit function in the sense $M(\boldsymbol{\Phi}^{(k+1)}) < M(\boldsymbol{\Phi}^{(k)})$. The ideal value $\alpha^{(k)} = 1$ may not always happen so that various modifications of the basic Newton method have to be implemented. The following basic model algorithm can be considered:

S1. If the conditions for convergence are satisfied, the algorithm terminates with $\boldsymbol{\Phi}^{(k)}$ as the solution.

S2. Compute a search direction $\Delta\boldsymbol{\Phi}^{(k)}$ solving (4.14).

S3. Compute the steplength $\alpha^{(k)} > 0$ for which $M(\boldsymbol{\Phi}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{\Phi}^{(k)}) < M(\boldsymbol{\Phi}^{(k)})$.

S4. Update the estimate for the minimum by $\boldsymbol{\Phi}^{(k+1)} := \boldsymbol{\Phi}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{\Phi}^{(k)}$, $k := k + 1$, and go back to step S1.

A standard convergence monitor in nonlinear programming is to choose the Euclidean norm $\|\boldsymbol{F}^{(\rho)}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\|$ of the residual produced by the KKT conditions (4.13) as a merit function. However, in many practical implementations, this choice of the merit function is not sufficient, since it does not allow to tell the difference between a local minimizer and a stationary non-minimizing point. The KKT conditions are necessary optimality conditions and hence, the optimization problem (4.1)–(4.2) and the nonlinear problem (4.13) are not equivalent, i.e., the Newton method may find solutions of (4.13) which do not minimize the objective function $f(\boldsymbol{x})$. Therefore, in order to find simultaneously solutions of both problems, a better approach is to rely on a hierarchy of two merit functions (cf., e.g., [GOW98, HPS02]). In general, the choice of merit functions in nonlinear constrained optimization problems is complicated. Several ideas have recently been proposed in the context of primal-dual interior methods (cf., e.g., [BHN99, ET*86, FGW02]). In particular, our *primary merit function* is related to those suggested in [GOW98] and is chosen as a modified augmented Lagrangian incorporating the logarithmic barrier function (4.8) as follows

$$M := M(\boldsymbol{x}, \boldsymbol{y}, \rho, \rho_A) = f(\boldsymbol{x}) - \rho \sum_{i=1}^{l} \log\, g_i(x) + \boldsymbol{y}^T \boldsymbol{h}(\boldsymbol{x}) + \frac{1}{2}\rho_A\, \boldsymbol{h}(\boldsymbol{x})^T \boldsymbol{h}(\boldsymbol{x}), \tag{4.18}$$

where ρ_A is a positive parameter. Our purpose now is to satisfy the descent conditions and to guarantee a reduction of the merit function in the sense that

each iterate should be an improved estimate of the solution of (4.6)–(4.7). Note that a descent is sought only with respect to $\boldsymbol{x}$ taking into account the original optimization problem. A standard way to achieve $M(\boldsymbol{x}+\alpha\Delta\boldsymbol{x},\boldsymbol{y},\rho,\rho_A) < M(\boldsymbol{x},\boldsymbol{y},\rho,\rho_A)$ is to require that $\Delta\boldsymbol{x}$ is a descent direction, i.e., $\Delta\boldsymbol{x}^T\nabla_{\boldsymbol{x}}M<0$, where $\nabla_{\boldsymbol{x}}M$ is the gradient of the primary merit function with respect to the primal variable $\boldsymbol{x}$. In particular, we have

$$\begin{aligned}\Delta\boldsymbol{x}^T\nabla_{\boldsymbol{x}}M &= \Delta\boldsymbol{x}^T(\boldsymbol{t}-\rho J_{\text{in}}^T G^{-1}\bar{\boldsymbol{e}}+J_{\text{eq}}^T\boldsymbol{y}+\rho_A J_{\text{eq}}^T\boldsymbol{h})\\ &= \Delta\boldsymbol{x}^T(\boldsymbol{t}-\rho J_{\text{in}}^T G^{-1}\bar{\boldsymbol{e}})-\boldsymbol{h}^T\boldsymbol{y}-\rho_A\boldsymbol{h}^T\boldsymbol{h},\end{aligned}\tag{4.19}$$

due to $J_{\text{eq}}\Delta\boldsymbol{x}=-\boldsymbol{h}$ from the second equation of (4.17). Hence, $\Delta\boldsymbol{x}^T\nabla_{\boldsymbol{x}}M<0$ can be satisfied if the augmented Lagrangian parameter ρ_A is sufficiently large, namely

$$\rho_A > \frac{\Delta\boldsymbol{x}^T(\boldsymbol{t}-\rho J_{\text{in}}^T G^{-1}\bar{\boldsymbol{e}})-\boldsymbol{h}^T\boldsymbol{y}}{\boldsymbol{h}^T\boldsymbol{h}}.$$

Hence, ρ_A can be changed within the optimization loop, if $\Delta\boldsymbol{x}$ is not a descent direction. In our algorithm, we choose

$$\rho_A = \min\left(\frac{5}{\boldsymbol{h}^T\boldsymbol{h}}(\Delta\boldsymbol{x}^T(\boldsymbol{t}-\rho J_{\text{in}}^T G^{-1}\bar{\boldsymbol{e}})-\boldsymbol{h}^T\boldsymbol{y}),100\right)\tag{4.20}$$

in the case $\Delta\boldsymbol{x}^T\nabla_{\boldsymbol{x}}M\geq 0$ and continue the loop (see [GOW98, HPS02] for details).

For the *secondary merit function* we choose the l_2- norm of the residual with respect to perturbed KKT-conditions (4.13). We apply the Newton method and choose the steplengths to strictly satisfy the inequality constraints $\boldsymbol{g}(\boldsymbol{x})>0$ and the complementarity constraints $\boldsymbol{z}>0$. Hence, the first requirement for the line-search approach is to insure a strict feasibility. Let $\hat{\alpha}$ and $\hat{\gamma}$ be separate steplengths defined as follows

$$\hat{\alpha}=\max\{\alpha|\boldsymbol{g}(\boldsymbol{x})+\alpha J_{\text{in}}\Delta\boldsymbol{x}\geq 0\},\qquad \hat{\gamma}=\max\{\gamma|\boldsymbol{z}+\gamma\Delta\boldsymbol{z}\geq 0\}.$$

Since we maintain interior (i.e., strict feasible) iterates, usually we take a parameter $\tau\in(0,1)$ bounded strongly away from unity and define $\alpha=\min(1,\tau\hat{\alpha})$ and $\gamma=\min(1,\tau\hat{\gamma})$. We use the same steplength γ for the Lagrange multiplier $\boldsymbol{y}$. In practice, both merit functions are used by means of the following strategy: If the steplengths α and γ lead to a reduction of M, they are accepted. If M does not decrease, we check the secondary merit function. If the latter decreases, the steplengths are accepted; otherwise damp the Newton steps by a certain factor and continue the procedure. The barrier parameter $\rho>0$ is updated by decreasing values until an approximate solution of the nonlinear problem is obtained (cf., e.g., [ET*86, GOW98, HPS02]). We rely on a *watchdog strategy* (see [CL*82]) to ensure progress in finding a local minimizer. If after some fixed number of iterations there is no reduction of M, the augmented Lagrangian parameter ρ_A is chosen sufficiently large in accordance with (4.20).

5 Solving the condensed primal–dual system

The discretized constrained optimization problem (3.2)–(3.3) is solved by the primal-dual interior-point method described in Sect. 4. We consider the diagonal matrices $D_1 := \mathrm{diag}(\alpha_i - \alpha_{\min})$ and $D_2 := \mathrm{diag}(\alpha_{\max} - \alpha_i)$ and introduce $\boldsymbol{z} := \rho D_1^{-1} \bar{\boldsymbol{e}} \geq 0$ and $\boldsymbol{w} = \rho D_2^{-1} \bar{\boldsymbol{e}} \geq 0$ serving as perturbed complementarity. We note that $1 \leq i \leq N$ where N is the number of finite elements in the discretized domain and $\bar{\boldsymbol{e}} = (1, 1, \ldots, 1)^T \in \mathcal{R}^M$. The primal-dual Newton-type interior-point method is applied to three sets of variables: primal feasibility $(\boldsymbol{u}, \boldsymbol{\alpha})$, dual feasibility $(\boldsymbol{\lambda}, \boldsymbol{\eta})$, and perturbed complementarity related to $(\boldsymbol{z}, \boldsymbol{w})$.

Denote the Lagrangian function of (3.2)–(3.3) by

$$\begin{aligned} \mathcal{L}(\boldsymbol{u}, \boldsymbol{\alpha}; \boldsymbol{\lambda}, \eta; \boldsymbol{z}, \boldsymbol{w}) := & \; f(\boldsymbol{u}, \boldsymbol{\alpha}) \\ & + \; \boldsymbol{\lambda}^T (A(\boldsymbol{\alpha})\, \boldsymbol{u} - \boldsymbol{f}) + \eta\, (g(\boldsymbol{\alpha}) - C) \\ & - \; \boldsymbol{z}^T (\boldsymbol{\alpha} - \alpha_{\min}\, \bar{\boldsymbol{e}}) - \boldsymbol{w}^T (\alpha_{\max} \bar{\boldsymbol{e}} - \boldsymbol{\alpha}). \end{aligned} \tag{5.1}$$

The Newton method applied to the KKT conditions of (5.1) results in

$$\begin{pmatrix} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} & 0 & 0 & 0 \\ \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}} & \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}} & \mathcal{L}_{\boldsymbol{\alpha}\eta} & -I & I \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} & 0 & 0 & 0 & 0 \\ 0 & \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 & 0 & 0 & 0 \\ 0 & Z & 0 & 0 & D_1 & 0 \\ 0 & -W & 0 & 0 & 0 & D_2 \end{pmatrix} \begin{pmatrix} \triangle \boldsymbol{u} \\ \triangle \boldsymbol{\alpha} \\ \triangle \boldsymbol{\lambda} \\ \triangle \eta \\ \triangle \boldsymbol{z} \\ \triangle \boldsymbol{w} \end{pmatrix} = - \begin{pmatrix} \nabla_{\boldsymbol{u}} \mathcal{L} \\ \nabla_{\boldsymbol{\alpha}} \mathcal{L} \\ \nabla_{\boldsymbol{\lambda}} \mathcal{L} \\ \nabla_{\eta} \mathcal{L} \\ \nabla_{\boldsymbol{z}} \mathcal{L} \\ \nabla_{\boldsymbol{w}} \mathcal{L} \end{pmatrix}, \tag{5.2}$$

where I stands for the identity matrix, $Z = \mathrm{diag}(z_i)$ and $W = \mathrm{diag}(w_i)$ are diagonal matrices. Following Sect. 4 we eliminate the increments for $\boldsymbol{z}$ and $\boldsymbol{w}$ from the 5th and 6th rows of (5.2), namely,

$$\triangle \boldsymbol{z} = D_1^{-1} (-\nabla_{\boldsymbol{z}} \mathcal{L} - Z \triangle \boldsymbol{\alpha}), \qquad \triangle \boldsymbol{w} = D_2^{-1} (-\nabla_{\boldsymbol{w}} \mathcal{L} + W \triangle \boldsymbol{\alpha}) \tag{5.3}$$

and substitute (5.3) in the second row of (5.2). We get the linear system $\tilde{K} \triangle \boldsymbol{\psi} = -\tilde{\boldsymbol{\xi}}$ for the increments of $\boldsymbol{\psi} := (\boldsymbol{u}, \boldsymbol{\alpha}, \boldsymbol{\lambda}, \eta)$, denoted by $\triangle \boldsymbol{\psi} := (\triangle \boldsymbol{u}, \triangle \boldsymbol{\alpha}, \triangle \boldsymbol{\lambda}, \triangle \eta)$ where $\tilde{K}$ is the matrix and $(-\tilde{\boldsymbol{\xi}})$ is the right-hand side of the following condensed primal-dual system

$$\begin{pmatrix} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} & 0 \\ \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}} & \tilde{\mathcal{L}}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}} & \mathcal{L}_{\boldsymbol{\alpha}\eta} \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} & 0 & 0 \\ 0 & \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 & 0 \end{pmatrix} \begin{pmatrix} \triangle \boldsymbol{u} \\ \triangle \boldsymbol{\alpha} \\ \triangle \boldsymbol{\lambda} \\ \triangle \eta \end{pmatrix} = - \begin{pmatrix} \nabla_{\boldsymbol{u}} \mathcal{L} \\ \tilde{\nabla}_{\boldsymbol{\alpha}} \mathcal{L} \\ \nabla_{\boldsymbol{\lambda}} \mathcal{L} \\ \nabla_{\eta} \mathcal{L} \end{pmatrix}. \tag{5.4}$$

The $\boldsymbol{\alpha}\boldsymbol{\alpha}$-entry of $\tilde{K}$ and the modified entry for the right-hand side are

$$\tilde{\mathcal{L}}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} + D_1^{-1}\, Z + D_2^{-1}\, W, \qquad \tilde{\nabla}_{\boldsymbol{\alpha}} \mathcal{L} = \nabla_{\boldsymbol{\alpha}} \mathcal{L} + D_1^{-1}\, \nabla_{\boldsymbol{z}} \mathcal{L} - D_2^{-1}\, \nabla_{\boldsymbol{w}} \mathcal{L}.$$

Direct methods for the solution of (5.4) can be divided into two classes: *range space methods* and *null space methods*. These approaches essentially

differ in the grouping of the matrix into a 2×2-block structure. The decomposition of the condensed primal-dual system (5.4) is related to the first approach. In this section, we consider the *null space* decomposition of the condensed primal-dual matrix interchanging the second and the third rows and columns. The resulting matrix can be written according to

$$\tilde{K} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} & 0 \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & 0 & \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} & 0 \\ \hline \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}} & \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}} & \tilde{\mathcal{L}}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathcal{L}_{\boldsymbol{\alpha}\eta} \\ 0 & 0 & \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 \end{array}\right),$$

where the first diagonal block

$$A_{11} = \begin{pmatrix} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & 0 \end{pmatrix} \tag{5.5}$$

is now an indefinite but nonsingular matrix. We remind that $\mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} = A(\boldsymbol{\alpha})$ is exactly the stiffness matrix corresponding to the equilibrium equation (2.12). Hence, A_{11}^{-1} exists, and the Schur complement $S := A_{22} - A_{21}A_{11}^{-1}A_{12}$ is defined correctly.

We use the following regular splitting of $\tilde{K}$

$$K^L \tilde{K}^R = M_1 - M_2 \tag{5.6}$$

with left and right factors given below and reasonable matrices M_1 and $M_2 \sim 0$. For solving the system of the form $\tilde{K}\triangle\boldsymbol{\psi} = -\tilde{\boldsymbol{\xi}}$, starting with an initial guess for $\triangle\boldsymbol{\psi} := (\triangle\boldsymbol{u}, \triangle\boldsymbol{\lambda}, \triangle\boldsymbol{\alpha}, \triangle\eta)^T$, the transforming iteration proposed in [Wit89] is described by

$$\triangle\boldsymbol{\psi}^{(\nu+1)} := \triangle\boldsymbol{\psi}^{(\nu)} + K^R M_1^{-1} K^L(-\tilde{\boldsymbol{\xi}} - \tilde{K}\triangle\boldsymbol{\psi}^{(\nu)}), \tag{5.7}$$

where the new iterate $\boldsymbol{\psi}^{(\text{new})}$ is obtained by a line-search in the direction $\triangle\boldsymbol{\psi}$, namely

$$\psi_j^{(\text{new})} = \psi_j^{(\text{old})} + \alpha_j(\triangle\boldsymbol{\psi})_j, \qquad 1 \le j \le 4.$$

The line-search approach and the choice of the steplengths parameters α_j are discussed in Sect. 4.3.

Using an appropriate preconditioner for the stiffness matrix, we approximate the first diagonal block (5.5) as follows

$$A_{11} = \begin{pmatrix} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & \tilde{\mathcal{L}}_{\boldsymbol{u}\boldsymbol{\lambda}} \\ \tilde{\mathcal{L}}_{\boldsymbol{\lambda}\boldsymbol{u}} & 0 \end{pmatrix} =: \tilde{A}_{11}.$$

Usually, the left and right transformations are of the form

$$K^L = I, \qquad K^R = \begin{pmatrix} I & -\tilde{A}_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 & -\tilde{\mathcal{L}}_{\boldsymbol{\lambda}\boldsymbol{u}}^{-1}\mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} & 0 \\ 0 & I & -\tilde{\mathcal{L}}_{\boldsymbol{u}\boldsymbol{\lambda}}^{-1}\mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix}.$$

In this case, the regular splitting (5.6) becomes $KK^R = M_1 - M_2$ where

$$M_1 = \begin{pmatrix} 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}} & 0 & 0 \\ \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}} & 0 & 0 & 0 \\ \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}} & \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}} & \tilde{S} & \mathcal{L}_{\boldsymbol{\alpha}\eta} \\ 0 & 0 & \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ R & Q \end{pmatrix} \tag{5.8}$$

and

$$M_2 = \begin{pmatrix} 0 & 0 & \mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} - \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}}\tilde{\mathcal{L}}^{-1}_{\boldsymbol{u}\boldsymbol{\lambda}}\mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}} & 0 \\ 0 & 0 & \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} - \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}}\tilde{\mathcal{L}}^{-1}_{\boldsymbol{\lambda}\boldsymbol{u}}\mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.9}$$

Note that $M_2 \sim 0$ if we have a good preconditioner for the stiffness matrix. In our numerical experiments, we choose a Cholesky decomposition of $\mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}}$. The second diagonal block Q in (5.8) is symmetric and indefinite given by

$$Q := \begin{pmatrix} \tilde{S} & \mathcal{L}_{\boldsymbol{\alpha}\eta} \\ \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 \end{pmatrix} \qquad \text{with} \qquad \tilde{S} := \tilde{\mathcal{L}}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} - \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}}\tilde{\mathcal{L}}^{-1}_{\boldsymbol{\lambda}\boldsymbol{u}}\mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}} - \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}}\tilde{\mathcal{L}}^{-1}_{\boldsymbol{u}\boldsymbol{\lambda}}\mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}}.$$

We denote the defect in (5.7) by $\boldsymbol{d} = -\tilde{\boldsymbol{\xi}} - \tilde{K}\triangle\boldsymbol{\psi}^{(\nu)}$ and compute the corresponding entries

$$\begin{aligned} \boldsymbol{d}_{\boldsymbol{u}} &= -\nabla_{\boldsymbol{u}}\mathcal{L} - \mathcal{L}_{\boldsymbol{u}\boldsymbol{\lambda}}\triangle\boldsymbol{\lambda} - \mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}}\triangle\boldsymbol{\alpha}, \\ \boldsymbol{d}_{\boldsymbol{\lambda}} &= -\nabla_{\boldsymbol{\lambda}}\mathcal{L} - \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{u}}\triangle\boldsymbol{u} - \mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}}\triangle\boldsymbol{\alpha}, \\ \boldsymbol{d}_{\boldsymbol{\alpha}} &= -\tilde{\nabla}_{\boldsymbol{\alpha}}\mathcal{L} - \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{u}}\triangle\boldsymbol{u} - \mathcal{L}_{\boldsymbol{\alpha}\boldsymbol{\lambda}}\triangle\boldsymbol{\lambda} - \tilde{\mathcal{L}}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}\triangle\boldsymbol{\alpha} - \mathcal{L}_{\boldsymbol{\alpha}\eta}\triangle\eta, \\ \boldsymbol{d}_{\eta} &= -\nabla_{\eta}\mathcal{L} - \mathcal{L}_{\eta\boldsymbol{\alpha}}\triangle\boldsymbol{\alpha}. \end{aligned}$$

Taking into account (5.7) one needs to compute $\boldsymbol{\delta} = M_1^{-1}\boldsymbol{d}$, i.e., $M_1\boldsymbol{\delta} = \boldsymbol{d}$. Consequently, we find $\delta_{\boldsymbol{\lambda}} = \tilde{\mathcal{L}}^{-1}_{\boldsymbol{u}\boldsymbol{\lambda}}\boldsymbol{d}_{\boldsymbol{u}}$ and $\delta_{\boldsymbol{u}} = \tilde{\mathcal{L}}^{-1}_{\lambda\boldsymbol{u}}\boldsymbol{d}_{\boldsymbol{\lambda}}$. To compute the remaining components of $\boldsymbol{\delta}$ we have to solve systems with an indefinite matrix Q of the form

$$\begin{pmatrix} \tilde{S} & \mathcal{L}_{\boldsymbol{\alpha}\eta} \\ \mathcal{L}_{\eta\boldsymbol{\alpha}} & 0 \end{pmatrix} \begin{pmatrix} \delta_{\boldsymbol{\alpha}} \\ \delta_{\eta} \end{pmatrix} = \begin{pmatrix} d_{\boldsymbol{\alpha}} - L_{\boldsymbol{\alpha}\boldsymbol{u}}\boldsymbol{\delta}_{\boldsymbol{u}} - L_{\boldsymbol{\alpha}\boldsymbol{\lambda}}\boldsymbol{\delta}_{\boldsymbol{\lambda}} \\ \boldsymbol{d}_{\eta} \end{pmatrix}.$$

Iterative procedures such as MINRES or Bi-CGSTAB (see [VdV92]) with appropriate stopping criteria can be applied in this case. Compute $K^R\boldsymbol{\delta}$ and find the increments from (5.7) as follows

$$\begin{aligned} \triangle\boldsymbol{u}^{(\text{new})} &= \triangle\boldsymbol{u}^{(\text{old})} + \delta_{\boldsymbol{u}} - \tilde{\mathcal{L}}^{-1}_{\boldsymbol{\lambda}\boldsymbol{u}}\mathcal{L}_{\boldsymbol{\lambda}\boldsymbol{\alpha}}\boldsymbol{\delta}_{\boldsymbol{\alpha}}, \quad \triangle\boldsymbol{\alpha}^{(\text{new})} = \triangle\boldsymbol{\alpha}^{(\text{old})} + \delta_{\boldsymbol{\alpha}}, \\ \triangle\boldsymbol{\lambda}^{(\text{new})} &= \triangle\boldsymbol{\lambda}^{(\text{old})} + \delta_{\boldsymbol{\lambda}} - \tilde{\mathcal{L}}^{-1}_{\boldsymbol{u}\boldsymbol{\lambda}}\mathcal{L}_{\boldsymbol{u}\boldsymbol{\alpha}}\boldsymbol{\delta}_{\boldsymbol{\alpha}}, \quad \triangle\eta^{(\text{new})} = \triangle\eta^{(\text{old})} + \delta_{\eta}. \end{aligned}$$

We apply the above algorithm (with a fixed number of iterations) to find the increments of the primal and dual variables $\triangle\boldsymbol{u}, \triangle\boldsymbol{\alpha}, \triangle\boldsymbol{\lambda}, \triangle\eta$ and then use (5.3) to determine the global search direction $\triangle\boldsymbol{\Phi}$.

6 Adaptive grid refinement

Advanced finite element applications in science and engineering provoke the extensive use of adaptive mesh-refinement techniques to optimize the number of degrees of freedom and obtain accurate enough numerical solutions. The adaptive framework requires a locally refined discretization in regions where a better accuracy is necessary.

The computation of the homogenized elasticity coefficients requires the numerical solution of (2.8) with the unit cell as the computational domain. Previous works on shape and topology optimization (cf., e.g., [Ben95, BS03, SZ92]) strongly suggest the use of locally refined grids particularly at material interfaces. In the context of shape optimization such local refinements have been mostly done before the computations relying on a priori geometric informations or in an interactive way (manual remeshing based on computational results). In case of local singularities of the discrete solution, a priori error estimates typically give information about the asymptotic error behavior and thus, are not the best choice to control the mesh. In those parts of the domain where the solution changes rapidly, an automatic grid refinement on the basis of reliable and robust a posteriori error estimators is highly beneficial. In practice, the main goal in adaptive mesh-refinement procedures is to refine the mesh so that the discretization error is within the prescribed tolerance and as possible equidistributed throughout the domain.

In the past twenty years, numerous studies have been devoted to an error control and a mesh-design based on efficient postprocessing procedures (cf., e.g., [CF01, EE*95, HP04a, ZZ87]). A natural requirement for a posteriori error estimates is to be less expensive than the cost of the numerically computed solution. Moreover, appropriate refinement techniques have to be applied to construct the adaptive mesh and implement the adaptive solver. Local reconstruction of the grid is necessary to be done with a computational cost proportional to the number of modified elements.

The a posteriori adaptive strategy can be described as follows:

A1. Start with an initial coarse mesh $\mathcal{T}_0$ fitting the domain geometry. Set $n := 0$.
A2. Compute the discrete solution on $\mathcal{T}_n$.
A3. Use a posteriori error indicator for each element $T \in \mathcal{T}_n$.
A4. If the global error is small enough, then stop. Else refine the marked elements, construct the next mesh $\mathcal{T}_{n+1}$, set $n := n+1$, and go to step A2.

The solution of our linear elasticity equation (2.8) is computed by using adaptive finite element method based on the *Zienkiewicz-Zhu* (referred as ZZ) error estimator. For instance, a recovery technique is analyzed in [ZZ87] for determining the derivatives (stresses) of the finite element solutions at nodes. The main idea of the recovery technique is to develop smoothing procedures which recover more accurate nodal values of derivatives from the original finite element solution.

The necessity of derivative recovering arises from the fact that in the finite element approach the rate of convergence of the derivatives is usually one order less than that of the discrete solution. In particular, the accuracy of the derivatives (stresses) computed by directly differentiating the discrete solution is inferior. Therefore, in many practical problems an improved accuracy of the stresses at nodes is needed.

Denote by $\boldsymbol{\sigma}$ the exact stress, by $\hat{\boldsymbol{\sigma}}$ the discrete finite element discontinuous stress, and by $\boldsymbol{\sigma}^*$ the smoothed continuous *recovered stress*. The computation of $\boldsymbol{\sigma}^*$ was proposed and discussed in [ZZ87] under the assumption that the same basis functions for interpolation of stresses are used as those for the displacements. The recovered stress $\boldsymbol{\sigma}^*$ is computed by smoothing the discontinuous (over the elements) numerical stress $\hat{\boldsymbol{\sigma}}$. The smoothing procedure can be accomplished by nodal averaging method or the L_2-projection technique. Note that the components of $\boldsymbol{\sigma}^*$ are piecewise linear and continuous.

The computational of the global L_2-projection is expensive and the authors of [ZZ87] proposed to use a lumping form of the mass matrix. Thus, the value of the recovered stress $\boldsymbol{\sigma}^*$ at a node P can be computed by averaging the stresses $\hat{\boldsymbol{\sigma}}$ at the elements that share that node. Denote by $Y_P \subset Y$ the neighborhood patch as an union of all triangles/tetrahedra T having node P. Consider

$$\boldsymbol{\sigma}^*(P) = \sum_{T \in Y_P} \omega|_T \, \hat{\boldsymbol{\sigma}}|_T, \quad \omega|_T = \frac{|T|}{|Y_P|}, \; T \in Y_P, \tag{6.1}$$

i.e., $\boldsymbol{\sigma}^*(P)$ is a weighted average of $\hat{\boldsymbol{\sigma}}$ with weights $\omega|_T$ defined on the elements belonging to Y_P. Least-square technique can also be applied to approximate the stress field at a given node.

It was shown in [ZZ87] that $\boldsymbol{\sigma}^*$ is a better approximation to $\boldsymbol{\sigma}$ than $\hat{\boldsymbol{\sigma}}$ and the following estimate holds

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^*\|_{0,Y} \ll \|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}\|_{0,Y}, \tag{6.2}$$

where Y is the periodicity microcell into consideration. Furthermore, the recovered technique was used in a formulation of a posteriori error estimator by comparing the recovered solution $\boldsymbol{\sigma}^*$ with the finite element solution $\hat{\boldsymbol{\sigma}}$. In particular, the estimate (6.2) allows us to replace the exact (unknown) stress $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^*$ and consider $\|\boldsymbol{\sigma}^* - \hat{\boldsymbol{\sigma}}\|_{0,Y}$ as an error estimator.

In many practical implementations reliability and efficiency are highly desirable properties in a posteriori error estimation. It basically means that there exist constants independent of the discrete solution and the mesh which limit the error (in a suitable norm) from below and above. Moreover, technically it is better to use local error estimators which are computationally less expensive. The following local estimator is considered

$$\eta_T := \|\boldsymbol{\sigma}^* - \hat{\boldsymbol{\sigma}}\|_{0,T}. \tag{6.3}$$

The nodal values of the recovered stresses are found locally. The elementwise contributions (6.3) are used further as local error indicators in the adaptive mesh-refinement procedure.

The global ZZ-error estimator is defined by

$$\eta_Y := \Big(\sum\nolimits_{T \in \mathcal{T}_n} \eta_T^2 \Big)^{1/2}. \tag{6.4}$$

Based on a posteriori processing, the local estimator (6.3) is practically efficient providing recovered values are more accurate, i.e., the quality of the a posteriori error estimator strongly depends on the approximation properties and the accuracy of the recovered solution.

Arbitrary averaging techniques in low order finite element applications for elasticity problems are subject of investigations in [CF01]. In the latter study the authors considered the following global averaging estimator

$$\eta_A := \min_{\boldsymbol{\sigma}^*} \ \|\boldsymbol{\sigma}^* - \hat{\boldsymbol{\sigma}}\|_{0,Y} \tag{6.5}$$

and proved an equivalence to the error $\|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}\|_{0,Y}$ with lower and upper bounds independent of the shape-regular mesh. Note that in (6.5) $\boldsymbol{\sigma}^*$ is a smoother approximation to $\hat{\boldsymbol{\sigma}}$ obtained by any averaging procedure. In particular, the final error estimate in [CF01] explains the reliability and robustness of the ZZ- a posteriori error estimators in practice.

7 Iterative solution techniques

In this section, we comment on the iterative solvers for the microcell problem (2.8) defined in Y to find the effective coefficients and for the homogenized elasticity equation (2.12) on the global domain Ω. After finite element discretization of the corresponding domain we get the following system of linear equations

$$A\,\boldsymbol{u} = \boldsymbol{f}, \tag{7.1}$$

where $\boldsymbol{u}$ is the vector of unknown displacements and $\boldsymbol{f}$ is the discrete right-hand side. The stiffness matrix A is symmetric and positive definite but not an M-matrix.

Two typical orderings of the unknowns are often used in practice. In the 3-dimensional case they are presented as follows

$$\left(u_1^{(x)}, u_1^{(y)}, u_1^{(z)}, u_2^{(x)}, u_2^{(y)}, u_2^{(z)}, \ldots, u_N^{(x)}, u_N^{(y)}, u_N^{(z)}\right), \tag{7.2}$$

referred to as a *pointwise displacements ordering* and

$$\left(u_1^{(x)}, u_2^{(x)}, \ldots, u_N^{(x)}, u_1^{(y)}, u_2^{(y)}, \ldots, u_N^{(y)}, u_1^{(z)}, u_2^{(z)}, \ldots, u_N^{(z)}\right), \tag{7.3}$$

called the *separate displacements ordering*. Here, $u_k^{(x)}$, $u_k^{(y)}$, and $u_k^{(z)}$ are the corresponding x, y-, and z- displacement components. For the the first ordering (7.2), the resulting stiffness matrix $A = A^{(point)}$ can be seen as a

discretization matrix consisting of elements which are small 3×3 blocks. For the second ordering (7.3), the matrix $A = A^{(block)}$ admits the following 3×3 block decomposition

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}. \tag{7.4}$$

In case of isotropic materials, the diagonal blocks $A_{jj},\ j = 1, 2, 3$, in (7.4) are discrete analogs of the following anisotropic Laplacian operators

$$\tilde{D}_1 = a\frac{\partial^2}{\partial x^2} + b\frac{\partial^2}{\partial y^2} + b\frac{\partial^2}{\partial z^2},\ \tilde{D}_2 = b\frac{\partial^2}{\partial x^2} + a\frac{\partial^2}{\partial y^2} + b\frac{\partial^2}{\partial z^2},\ \tilde{D}_3 = b\frac{\partial^2}{\partial x^2} + b\frac{\partial^2}{\partial y^2} + a\frac{\partial^2}{\partial z^2}$$

with coefficients $a = E(1-\nu)/((1+\nu)(1-2\nu))$ and $b = 0.5E/(1+\nu)$ where E is the Young modulus and ν is the Poisson ratio of the corresponding material. This anisotropy requires a special care to construct an efficient preconditioner for the iterative solution method. Based on Korn's inequality, it can be shown that A and its block diagonal part are spectrally equivalent. The condition number of the preconditioned system depends on the Poisson ratio ν of the materials and the constant in the Korn inequality. For the background of the spectral equivalence approach using block diagonal displacement decomposition preconditioners in linear elasticity problems we refer to [BLA94]. Note that the spectral equivalence estimate will deteriorate for ν close to 0.5 which is not the case in our particular applications.

The PCG method is applied to solve the linear system (7.1). We propose two approaches to construct a preconditioner for A:

(i) construct a preconditioner for $A^{(point)}$

(ii) construct a preconditioner for $A^{(block)}$ of the type $M = \mathrm{diag}(M_{jj})$, where $M_{jj} \sim A_{jj},\ j = 1, 2, 3$, are "good" approximations to the diagonal blocks of A. In case (i) we have chosen the incomplete Cholesky (IC) factorization of A with an appropriate stopping criterion.

An efficient preconditioner for A_{jj} in case (ii) turns out to be a matrix M_{jj} corresponding to a Laplacian operator $(-\mathrm{div}\,(c\,\mathrm{grad}\,u))$ with a fixed scale factor c. In our case we use, for instance, $c = b/2$ for all three diagonal blocks. *Algebraic MultiGrid* (AMG) method is applied as a "plug-in" solver for A (see [RS86] for details). This method is a purely matrix-based version of the algebraic multilevel approach and has shown in the last decade numerous efficient implementations in solving large sparse unstructured linear systems of equations without any geometric background.

8 Numerical experiments

In this section, we comment on some computational results concerning the microscopic problem to find the homogenized elasticity coefficients and the macroscopic shape optimization problem. For simplicity, we suppose linear elasticity with homogeneous and isotropic constituents in terms of carbon

and SiC. The Young modulus E (in GPa) and the Poisson ratio ν of our two materials are, respectively, $E = 10,\ \nu = 0.22$ for carbon and $E = 410,\ \nu = 0.14$ for SiC.

The computation of the characteristic displacement fields $\boldsymbol{\xi}^{kl}$ and the homogenized elasticity coefficients (2.9) requires the solution of linear elastic boundary value problems with the periodicity cell Y as the computational domain. The elasticity equation (2.8) is solved numerically using a conforming finite element discretization of the periodicity cell Y by linear basis functions. Since the periodic displacements $\boldsymbol{\xi}^{kl} = \boldsymbol{\xi}^{lk}$ are symmetric, the equation (2.8) is computed numerically 3 times in the case $d = 2$ and respectively, 6 times in the case $d = 3$. Due to the composite character of our microcell there are material interfaces where the solution changes significantly. Hence, local refinement of the underlying finite element mesh is strongly advised. As discussed in Sect. 6, we use an adaptive grid refinement strategy based on a posteriori error estimator of Zienkiewicz-Zhu type obtained by local averaging of the computed stress tensor. Note that the adaptivity procedure is local and computationally cheap.

Denote the global density of the solid material part in the microstructure by $\mu,\ 0 < \mu < 1$. Note that the density of the tracheidal cells of the wood essentially depends on the growth of the tree. If μ is relatively small, we speak about an *early wood* (grown in spring and summer) and respectively, about *late wood* (grown in autumn and winter) for values of μ, close to 1.

We present first some numerical experiments on a plane microstructure ($d = 2$) shown in Fig. 2.2 b). More experiments can be found in [HP04a]. We assume that the material layers in the periodicity cell have equal widths from all sides of the cell. Denote by $\alpha_i,\ i = 1, 2$, the widths of the carbon and SiC layers, respectively. Figure 8.1 a) illustrates the behavior of the homogenized coefficient E^H_{1212} in case of square hole versus α_1 and α_2 which vary between 0 and 0.5. We compute the effective coefficients E^H_{ijkl} only for a fixed number of values of the design parameters (e.g., 20×20 grid as shown on Fig. 8.1) and then interpolate the values by splines. With regard to the homogenized state equation (2.12), this procedure results in having explicit formulas at hand for the gradients and the Hessian of the Lagrangian function needed in the optimization procedure.

In principal, the hole is located inside the microstructure but we find interesting to demonstrate the behavior, for instance, of E^H_{1212} depending on a rectangular hole $[1 - a] \times [1 - b]$, see Fig. 8.1 b). Note that $a = b = 0$ represents a complete void, $a = b = 1$ realizes a complete solid material, and $0 < a < 1,\ 0 < b < 1$ characterize a general porous material. We consider in this example the case when the carbon has completely reacted with the SiC which strongly concerns the so-called *pure biomorphic SiC-ceramics*. Very recently, the chemical experiments have shown that the carbon phase limits the mechanical properties of the composite materials and restricts their high-temperature applications. The final transformation of the original carbonized template to pure ceramic composite requires to offer enough silicon during

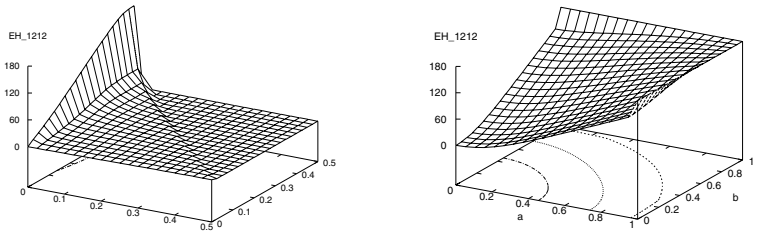

Fig. 8.1. Homogenized coefficient E^H_{1212}: a) w.r.t. the widths of carbon and SiC layers (square hole); b) w.r.t. the sizes $1-a$ and $1-b$ of the rectangular hole (pure SiC-ceramic)

the infiltration process and to wait an appropriate time until the carbon is completely consumed by the silicon resulting in a SiC-phase.

Figure 8.2 displays the dependence of the homogenized elasticity coefficients on the density μ of the cell. In particular, we show this behavior versus the width of the SiC layer in case of pure SiC-ceramics. Figure 8.2 a) shows the behavior of the effective coefficients for early wood ($0 \leq \alpha_2 \leq 0.15$, $\mu = 51\%$) and Fig. 8.2 b) demonstrates the coefficients for late wood ($0 \leq \alpha_2 \leq 0.3$, $\mu = 84\%$). One can easily observe from both pictures on this figure a highly nonlinear behavior of the homogenized coefficients.

The mesh-adaptive process is visualized in Fig. 8.3. We see that in case of one material available in the microstructure, an appropriate refinement is done around the corners where the hole with a complete pore is located, see Fig. 8.3 a). In case of more materials, additional mesh-adaptivity is needed across the material interfaces in the microstructure due to the strongly varying material properties in terms of Young's modulus and Poisson's ratio.

In Table 8.1 we give some results for the homogenized elasticity coefficients on the first ten adaptive refinement levels for various values of the density. We report the number of triangles NT and the number of nodes NN on each level when solving problem (2.8). We see from the computed values that the mesh sensitivity on the successive levels is very small. Our adaptive mesh-refinement

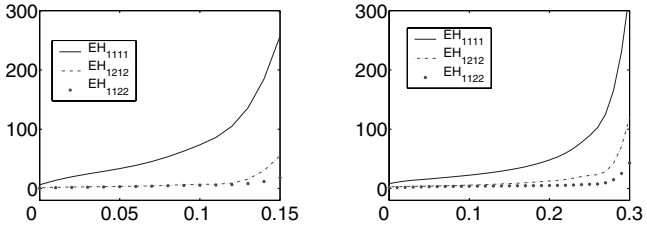

Fig. 8.2. Homogenized coefficients w.r.t. the width α_2 of SiC layer for pure SiC-ceramic: a) early wood, density 51%; b) late wood, density 84%

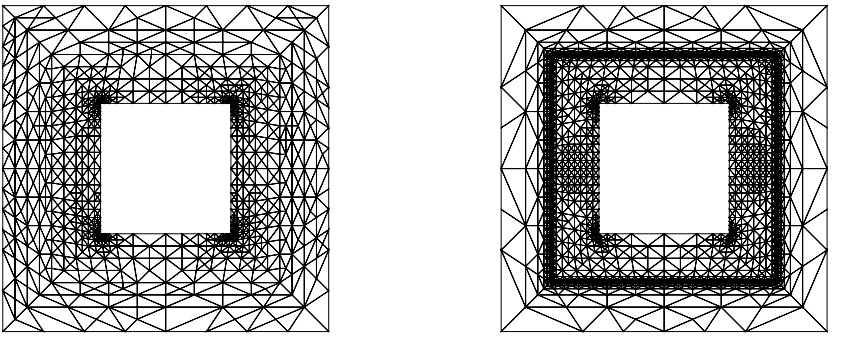

Fig. 8.3. Late wood, density $\mu = 84\%$, 9 adaptive refinement levels: a) SiC, 1527 triangles, 818 nodes; b) carbon and SiC, 3752 triangles, 1916 nodes

procedure stops when a priori given limit for the number of refinement levels is reached.

We are now concerned with the solution of problem (3.2)–(3.3). In Table 8.2 we report some numerical results from running the optimization code varying the constant C with respect to (3.3). Our purpose is to find the optimal widths/lengths of the layers in the composite material and to show the convergence behavior of the optimization algorithm. The domain Ω is chosen to be a circle which corresponds naturally to a cross section of the original wood structure. We have fixed the discretization and vary the initial values for the lengths of the carbon and SiC layers denoted, respectively, by $\alpha_1^{(0)}$ and $\alpha_2^{(0)}$. As before, we report the number of iterations ITER to get convergence, the optimal lengths α_1 and α_2 of the carbon and SiC layers, the last value of the barrier parameter ρ, the final value of the primary merit function M, the l_2-norm of the residual, and the l_2-norm of the complementarity conditions $\boldsymbol{v} = (\boldsymbol{z}, \boldsymbol{w})$ at the last iteration. We see from the experiments that the optimal

Table 8.1. Homogenized coefficients w.r.t. refinement level, a) $\mu = 51\%$, b) $\mu = 84\%$

level	E^H_{1111}	E^H_{1122}	E^H_{1212}	NT	NN
1	64.975	7.664	12.116	168	100
2	63.336	6.642	9.750	220	126
3	58.466	6.682	8.073	288	162
4	56.572	7.012	6.643	484	262
5	54.385	6.245	6.212	712	378
6	52.936	6.091	5.474	1208	630
7	51.914	5.458	5.306	1800	932
8	50.861	4.790	5.217	2809	1444
9	50.455	4.571	5.029	3754	1919
10	49.591	4.359	4.983	5918	3013

level	E^H_{1111}	E^H_{1122}	E^H_{1212}	NT	NN
1	33.430	3.885	9.893	168	100
2	33.064	3.929	9.577	216	126
3	32.844	4.024	9.283	300	168
4	32.291	4.254	8.970	544	296
5	32.144	4.312	8.809	828	438
6	31.909	4.372	8.703	1354	705
7	31.862	4.379	8.526	1892	980
8	31.735	4.399	8.470	2894	1485
9	31.711	4.400	8.373	3752	1916
10	31.487	4.497	8.321	5716	2906

Table 8.2. Convergence results for biomorphic microcellular SiC ceramics

$\alpha_1^{(0)}$	$\alpha_2^{(0)}$	C	ITER	α_1	α_2	ρ	M	$\|\boldsymbol{F}^{(\rho)}\|_2$	$\|\boldsymbol{v}\|_2$
0.05	0.05	0.3	11	3.6e-12	0.3	1.3e-17	1.24	9.63e-6	e-10
0.1	0.1	0.3	11	5.5e-14	0.3	3.0e-21	1.24	1.03e-6	e-12
0.1	0.1	0.4	12	1.6e-16	0.4	1.2e-26	0.85	8.63e-9	e-14
0.2	0.2	0.1	16	5.5e-17	0.1	2.2e-25	7.73	2.23e-8	e-13
0.2	0.2	0.2	13	1.0e-16	0.2	5.3e-26	2.34	1.54e-8	e-14
0.2	0.2	0.3	11	2.5e-16	0.3	6.7e-26	1.24	1.79e-8	e-14
0.24	0.24	0.15	11	5.4e-15	0.15	4.1e-12	3.81	4.99e-7	e-12
0.3	0.1	0.4	11	1.3e-12	0.4	8.5e-19	0.85	5.07e-6	e-10
0.4	0.05	0.1	17	9.8e-15	0.1	6.9e-21	7.73	9.49e-7	e-11

length α_1 of the carbon layer in all the runs is very close to zero, i.e., the solid part of the body is entirely occupied by a silicon carbide layer due to the higher stiffness of this material.

In case of 3-dimensional implementations we decompose the periodic microcell Y first in hexahedra and further we use continuous, piecewise linear finite elements on tetrahedral shape regular meshes. The adaptive grid refinement process is visualized in Fig. 8.4. The mesh adaptivity around the material interfaces has been realized by means of Zienkiewicz-Zhu type a posteriori error which is used heuristically (as an error indicator). One computes the error (6.3) locally for each element and mark for refinement those tetrahedra $\{T\}$ for which

$$\eta_T \geq \gamma \max_{T' \in \mathcal{T}_n} \eta_{T'},$$

where $0 < \gamma < 1$ is a prescribed threshold, for instance, $\gamma = 0.5$. The refinement process is visualized in Fig. 8.4 b) on the cross section of the microstruc-

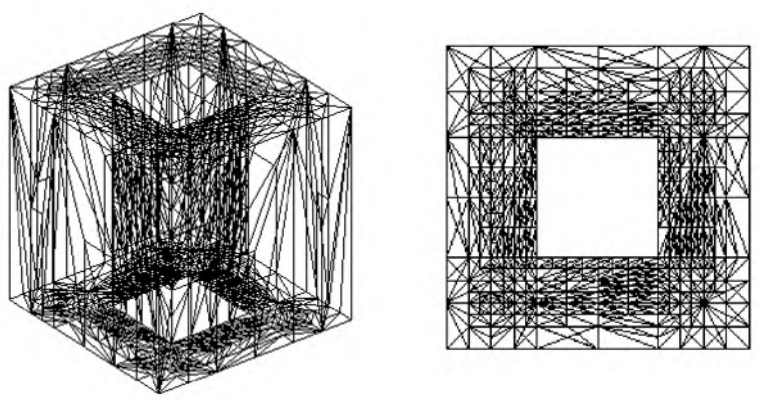

Fig. 8.4. Adaptive refinement a) 3-D unit periodicity cell Y, b) Cross section of Y

Table 8.3. Homogenized coefficients for late wood, density $\mu = 91\%$

level	E^H_{1111}	E^H_{2222}	E^H_{3333}	E^H_{1212}	E^H_{2323}	E^H_{1313}
1	148.35	152.57	153.96	60.22	62.46	59.50
2	154.34	162.64	162.77	69.71	71.31	65.79
3	142.66	148.42	162.79	60.51	65.26	63.23
4	145.84	137.61	161.70	53.91	59.04	62.92
5	127.99	134.32	161.43	49.41	56.19	56.49
6	98.29	111.65	160.71	40.44	46.14	48.45
7	91.79	90.23	158.29	35.70	43.69	46.03
8	82.42	83.00	160.57	30.59	41.03	43.70
9	75.05	75.11	160.22	26.93	39.75	40.97
10	69.66	70.30	159.82	25.47	37.16	39.30

ture Y for widths of the C- and SiC- layers $\alpha_1 = \alpha_2 = 0.15$. Additional adaptive refinement is generated in the stiffer material (SiC) and on the interface between the materials due to the different characteristic constants.

In Table 8.3 we report some values of the computed 3-dimensional homogenized coefficients with respect to the adaptive refinement level for a late wood with density $\mu = 91\%$. More numerical experiments for various values of the density and various number of adaptive levels can be found in [HP06a].

Table 8.4 presents some convergence results for the proposed preconditioners within PCG method. For various values of the density μ of the periodical microstructure we report the number of degrees of freedom NDOF, the number of iterations ITER, and the CPU-time in seconds for the first 11 adaptive refinement levels. One can see from the numerical results a better convergence of AMG-preconditioner compared to IC-factorization. We observe an essential efficiency of AMG for a larger number of unknowns.

Table 8.4. Convergence results with IC and AMG preconditioners, density μ

density	level	1	2	3	4	5	6	7	8	9	10	11
$\mu = 51\%$	NDOF	78	90	126	225	336	579	1185	1908	3360	5598	9987
IC	ITER	9	8	14	23	40	66	105	150	235	269	299
	CPU	e-16	e-16	e-16	0.1	0.2	0.2	0.9	2.4	8.2	20.9	59.1
AMG	ITER	11	13	13	15	18	23	38	57	89	94	99
	CPU	e-16	e-16	e-16	0.2	0.3	0.5	1.5	3	7.6	14.8	23.5
$\mu = 84\%$	NDOF	78	93	150	261	510	1047	2103	3843	6537	10485	18459
IC	ITER	10	11	16	21	44	78	117	171	226	273	301
	CPU	e-16	e-16	0.1	0.1	0.1	0.6	2.4	8.4	24.3	63.7	187.1
AMG	ITER	12	14	14	14	18	31	43	73	69	74	75
	CPU	e-16	e-16	e-16	0.2	0.4	1.1	3	7.5	15.5	25.6	33.8

Acknowlegdments. Research supported by DFG Priority Program 1095 "Analysis, Modeling and Simulation of Multiscale Problems" under Ho 877/5.

References

[All02] G. Allaire. Shape Optimization by the Homogenization Method. Springer, Berlin-Heidelberg-New York, 2002.

[BP84] N. Bakhvalov and G. Panasenko. Averaging Processes in Periodic Media. Nauka, Moscow, 1984.

[Ben95] M.P. Bendsøe. Optimization of Structural Topology, Shape, and Material. Springer, Berlin, 1995.

[BS03] M.P. Bendsøe and O. Sigmund. Topology Optimization: Theory, Methods and Applications. Springer, Berlin-Heidelberg-New York, 2003.

[BLP78] A. Bensoussan, J.L. Lions, and G. Papanicolaou. Asymptotic Analysis for Periodic Structures. Elsevier, North-Holland, Amsterdam, 1978.

[BLA94] R. Blaheta. Displacement decomposition – incomplete factorization preconditioning techniques for linear elasticity problems. *Numer. Linear Algebra Appl.*, 1(2), 107–128, 1994.

[BHN99] R.H. Byrd, M.E. Hribar, and J. Nocedal. An interior point algorithm for large scale nonlinear programming. *SIAM J. Optim.*, 9(4), 877–900, 1999.

[CF01] C. Carstensen and S. A. Funken. Averaging technique for FE–a posteriori error control in elasticity. Part I: Conforming FEM. *Comput. Methods Appl. Mech. Engrg.*, 190, 2483–2498, 2001.

[CL*82] R.M. Chamberlain, C. Lemaréchal, H.C. Pedersen, and M.J.D. Powell. The watchdog technique for forcing convergence in algorithms for constrained optimization. *Math. Progr. Study*, 16, 1–17, 1982.

[Eli00] M. Elices. Structural Biomaterials. Princeton University Press, 2000.

[ET*86] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya, and Y. Zhang. On the formulation and theory of the Newton interior–point method for nonlinear programming. *J. Optim. Theory Appl.*, 89, 507–541, 1996.

[EE*95] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations. *Acta Numerica*, 105–158, 1995.

[FM68] A.V. Fiacco and G.P. McCormick. Nonlinear Programming. Sequential Unconstrained Minimization Techniques. John Wiley and Sons, New York, New York, 1968. Republished by SIAM, Philadelphia, Pennsylvania, 1990.

[Fri55] K.R. Frisch. The Logarithmic Potential Method of Convex Programming. Memorandum, University Institute of Economics, Oslo, Norway, 1955.

[FGS96] A. Forsgren, P.E. Gill, and J.R. Shinnerl. Stability of symmetric ill–conditioned systems arising in interior methods for constrained optimization. *SIAM J. Matrix Anal. Appl.*, 17, 187–211, 1996.

[FGW02] A. Forsgren, P.E. Gill, and M.H. Wright. Interior methods for nonlinear optimization. *SIAM Review*, 44, 525–597, 2002.

[GOW98] D.M. Gay, M.L. Overton, and M.H. Wright. A primal–dual interior method for nonconvex nonlinear programming. *Advances in Nonlinear Programming*, (Y. Yuan, ed.), Kluwer, Dordrecht, Holland, 31–56, 1998.

[GA88] L.J. Gibson, M.F. Ashby. Cellular Solids, Structure, and Properties. Pergamon Press, New York, 1988.

[GLK98a] P. Greil, T. Lifka, and A. Kaindl. Biomorphic cellular silicon carbide ceramics from wood: I. Processing and microstructure. *J. Europ. Ceramic Soc.*, 18, 1961–1973, 1998.

[GLK98b] P. Greil, T. Lifka, and A. Kaindl. Biomorphic cellular silicon carbide ceramics from wood: II. Mechanical properties. *J. Europ. Ceramic Soc.*, 18, 1975–1983, 1998.

[HPS02] R.H.W. Hoppe, S.I. Petrova, and V. Schulz, Primal–dual Newton–type interior–point method for topology optimization. *J. Optim. Theory Appl.*, 114(3), 545–571, 2002.

[HP04a] R.H.W. Hoppe and S.I. Petrova. Optimal shape design in biomimetics based on homogenization and adaptivity. *Math. Comput. Simul.*, 65, 257–272, 2004.

[HP04b] R.H.W. Hoppe and S.I. Petrova. Primal–dual Newton interior point methods in shape and topology optimization. *Numer. Linear Algebra Appl.*, 11(5-6), 413–429, 2004.

[HP06a] R.H.W. Hoppe and S.I. Petrova. Efficient solvers for 3-D homogenized elasticity model. In J. Dongarra et al., eds., *Lect. Notes Comput. Sci.*, Springer, 3732, 857–863, 2006.

[HDL02] K.A. Hudgins, A.K. Dillow, and A.M. Lowman. Biomimetic Materials and Design: Biointerfacial Strategies, Tissue Engineering, and Targeted Drug Delivery. Marcel Dekker, New York, 2002.

[JKO94] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. Homogenization of Differential Operators and Integral Functionals. Springer, 1994.

[OT83] N. Olhoff and J.E. Taylor. On structural optimization. *J. Appl. Mech.*, 50, 1139–1151, 1983.

[OT*95] T. Ota, M. Takahashi, T. Hibi, M. Ozawa, S. Suzuki, Y. Hikichi, and H. Suzuki. Biomimetic process for producing SiC wood. *J. Amer. Ceram. Soc.*, 78, 3409–3411, 1995.

[RS86] J.W. Ruge and K. Stüben. Algebraic multigrid (AMG). In S.F. McCormick, ed., *Multigrid Methods, Frontiers in Applied Mathematics*, volume 5, SIAM, Philadelphia, 1986.

[SP80] E. Sanchez-Palencia, Non–homogeneous Media and Vibration Theory. Lecture Notes in Physics, volume 127, Springer, Berlin-Heidelberg, 1980.

[SZ92] J. Sokolowski and J.-P. Zolésio. Introduction to Shape Optimization. Springer Series in Computational Mathematics, volume 16, Springer, 1992.

[SK91] K. Suzuki and N. Kikuchi. A homogenization method for shape and topology optimization. *Comput. Meth. Appl. Mech. Engrg.*, 93, 291–318, 1991.

[VS99] R.J. Vanderbei and D.F. Shanno. An interior–point algorithm for nonconvex nonlinear programming. *Comput. Optim. Appl.*, 13, 231–252, 1999.

[VdV92] H.A. Van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13, 631–644, 1992.

[VSG02] E. Vogli, H. Sieber, P. Greil. Biomorphic SiC-ceramic prepared by Si-gas phase infiltration of wood. *J. Europ. Ceramic Soc.*, 22, 2663–2668, 2002.

[Wit89] G. Wittum. On the convergence of multigrid methods with transforming smoothers. Theory with applications to the Navier–Stokes equations. *Numer. Math.*, 57, 15–38, 1989.

[Wri98] M.H. Wright. Ill–conditioning and computational error in interior methods for nonlinear programming. *SIAM J. Optim.*, 9, 84–111, 1998.

[ZZ87] O.C. Zienkiewicz and J.Z. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *Intern. J. Numer. Methods Engrg.*, 24, 337–357, 1987.