Applications of Primal-dual Interior Methods in Structural Optimization

Ronald H. W. Hoppe, Svetozara I. Petrova

Angaben zur Veröffentlichung / Publication details:

Hoppe, Ronald H. W., and Svetozara I. Petrova. 2003. "Applications of Primal-dual Interior Methods in Structural Optimization." *Computational Methods in Applied Mathematics* 3 (1): 159–76. https://doi.org/10.2478/cmam-2003-0011.

COMPUTATIONAL METHODS IN APPLIED MATHEMATICS, Vol.3(2003), No.1, pp.159–176

APPLICATIONS OF PRIMAL-DUAL INTERIOR METHODS IN STRUCTURAL OPTIMIZATION

RONALD H.W. HOPPE

Institute of Mathematics, University of Augsburg
University Str. 14, D-86159 Augsburg, Germany
E-mail: hoppe@math.uni-augsburg.de

SVETOZARA I. PETROVA [1]

Institute of Mathematics, University of Augsburg
University Str. 14, D-86159 Augsburg, Germany
E-mail: petrova@math.uni-augsburg.de

Dedicated to Raytcho Lazarov on the occasion of his 60th birthday.

Abstract — We are concerned with structural optimization problems for technological processes in material science that are described by partial differential equations. In particular, we consider the topology optimization of conductive media in high-power electronic devices described by Maxwell equations and the optimal design of composite ceramic materials by homogenization modeling. All these tasks lead to constrained nonconvex minimization problems with both equality and inequality constraints on the state variables and design parameters. After discretization by finite elements, we solve the discretized optimization problems by a primal-dual Newton interior-point method. Within a line-search approach, transforming iterations are applied with respect to the null space decomposition of the condensed primal-dual system to find the search direction. Some numerical experiments for the two applications are presented.

2000 Mathematics Subject Classification: 65K10; 74B05; 74Q05; 90C30.

Keywords: structural optimization, nonlinear programming, primal-dual interior-point methods, eddy current equations, elasticity equations.

1. Introduction

Structural optimization has recently been the subject of much attention in various problems of optimal design and construction of structures (cf., e.g., [1,23] 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

[1]Permanent address: Central Laboratory for Parallel Processing, Bulgarian Academy of Sciences, Acad. G. Bontchev Str., Block 25A, 1113 Sofia, Bulgaria

are different types in structural optimization. Detailed classification of these problems is given, for instance, in [22]. 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. In the topology optimization of solid structures, we are interested in the determination of the optimal placement of a 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.

In all optimization algorithms, we are typically faced with constrained nonconvex minimization problems with both equality and inequality constraints on the state variables and design parameters. After discretization, for example, by finite elements, we solve the discretized optimization problems by primal-dual Newton interior-point methods, originally proposed for linear programming in [21]. The main idea of these methods is to generate iteratively approximations of the solution which strictly satisfy the inequality constraints. Recently, primal-dual interior methods have been extended to nonlinear programming problems (cf., e.g., [4, 5, 7–9, 11–13, 25]).

This paper is organized as follows. In Section 2, we formulate two structural optimization problems. The first one is related to the topology optimization in electromagnetic media described by Maxwell equations. In the stationary case, the mathematical model leads to elliptic boundary value problems corresponding to a minimization of the energy dissipation. In this model, the state variable is the associated electric potential and the design variable is the electric conductivity. We are interested in finding the optimal distribution of the conductivity in a fixed geometrical domain. Our second application concerns the shape optimization of recently produced biomorphic microcellular silicon carbide ceramics from wood (see [15]). Based on the constitutive microstructural model, the macroscale model is obtained by using the homogenization approach. In the past twenty years, the homogenization method has found a lot of important applications in mechanics of composite materials (cf., e.g., [18–20]). We comment on the optimal distribution of our composite material in a suitable reference domain which can carry given loads. The shape optimization problem is solved under a set of constraints on the state variables (displacements in terms of the elasticity equations) and design parameters (lengths of the layers in the microstructure, angle of cell rotation, etc.). Section 3 concerns general nonconvex nonlinear programming with equality and inequality constraints when the first and second derivatives of the objective and constraint functions are available. The damped Newton method is applied to the perturbed Karush-Kuhn-Tucker conditions for the logarithmic barrier function problem. A line-search approach is used to find the search direction. Section 4 focuses on the solver of the condensed primal-dual system. We use the null space decomposition of the condensed primal-dual matrix and apply the method of transforming iterations (see [28]) to find an appropriate solution for the search direction. In the last section, we give some numerical experiments applying the primal-dual interior-point method to the nonlinear problem and present the respective material distributions and convergence results for both applications.

2. Structural optimization problems

2.1. Maxwell equations

In this section, we consider an optimal design problem arising in high-power electronics, namely, in the design of converter modules that are used as electric drives for high-power electromotors. The electromagnetic fields in the low frequency regime can be described by the quasi-stationary limit of Maxwell equations also known as the *eddy current equations*

$$\frac{\partial B}{\partial t} + \mathrm{curl} E = 0, \quad \mathrm{div} B = 0, \quad \mathrm{curl} H = J, \tag{1}$$

$$B = \mu H, \quad J = \sigma E. \tag{2}$$

Here, E and H denote the electric and the magnetic field, B is the magnetic induction, J stands for the current density, and the material parameters μ and σ refer to the magnetic permeability and the electric conductivity, respectively.

We use a potential formulation by introducing a scalar electric potential u and a magnetic vector potential A according to

$$E = -\mathrm{grad}\, u - \frac{\partial A}{\partial t}, \quad B = \mathrm{curl} A$$

(cf., e.g., [3]). Then, (1)–(2) give rise to the following coupled system of PDEs for the electromagnetic potentials u and A

$$\mathrm{div}(\sigma \mathrm{grad}\, u) = 0 \quad \text{in} \quad \Omega, \tag{3}$$

$$\sigma n \cdot \mathrm{grad}\, u = \begin{cases} I_\nu & \text{on} \quad \Gamma_\nu \subset \partial\Omega, \\ 0 & \text{elsewhere}, \end{cases} \tag{4}$$

$$\sigma \frac{\partial A}{\partial t} + \mathrm{curl}\mu^{-1}\mathrm{curl} A = \begin{cases} -\sigma \mathrm{grad}\, u & \text{in} \quad \Omega, \\ 0 & \text{in} \quad R^3 \setminus \bar{\Omega} \end{cases} \tag{5}$$

the latter one with appropriate initial and boundary conditions. Note that in (4) we refer to I_ν as the fluxes associated with the contacts $\Gamma_\nu \subset \partial\Omega, 1 \leqslant \nu \leqslant N_c$, satisfying the compatibility conditions $\sum_{\nu=1}^{N_c} I_\nu = 0$.

The electric energy dissipation given by the Joule-Lenz law reads as follows:

$$f(u, \sigma, A) = \int\limits_\Omega J \cdot E dx.$$

In particular, in the stationary regime this reduces to

$$f(u, \sigma) = -\int\limits_\Omega J \cdot \mathrm{grad} u\, dx = -\int\limits_\Omega \mathrm{div}(uJ)\, dx. \tag{6}$$

The last equality in (6) follows from $\mathrm{div}(uJ) = J \cdot \mathrm{grad} u + u\, \mathrm{div} J$, taking into account that $\mathrm{div}\, J = 0$ in view of (1). Using the Gauss theorem and the Neumann boundary conditions, from (4) we get

$$f(u, \sigma) = -\int\limits_{\partial\Omega} n \cdot J\, ds u = \sum_{\nu=1}^{N_c} \int\limits_{\Gamma_\nu} I_\nu u ds. \tag{7}$$

We solve the optimization problem for the energy dissipation given by (7). Find

$$\min_{u,\sigma} f(u,\sigma) \;=\; \min_{u,\sigma} \sum_{\nu} \int\limits_{\Gamma_\nu} I_\nu u \; ds, \tag{8}$$

subject to the following constraints

$$A(\sigma)\, u - b = 0, \quad g(\sigma) - C = 0, \quad \sigma - \sigma_{\min} e \geqslant 0, \quad \sigma_{\max} e - \sigma \geqslant 0, \tag{9}$$

where A is the stiffness matrix related to the electric potential equation (3), b is the discrete load vector, $e = (1, 1, \ldots, 1)^T$, σ is the discrete conductivity function, and $g(\sigma) = \int\limits_{\Omega} \sigma \, dx$. Here, $\sigma_{\min}$ and $\sigma_{\max}$ are a priori given positive limits for the conductivity, and C is a fixed given value in the mass constraint. Note that in this case we take $\sigma_{\min} > 0$ in order to keep the ellipticity of the discrete problem.

2.2. Elasticity equations

In this section, we consider the recently produced biomorphic microcellular silicon carbide (SiC) ceramics obtained from wood (for a detailed description of the production and processing scheme, see [15]). We are interested in the optimal performance of the new composite materials which can be obtained by tuning the microstructural geometrical features that strongly influence the macrocharacteristics of the final products. Our macroscale model is obtained by using the homogenization approach which has found a lot of practical applications in structural mechanics to determine the optimal design of microstructured materials (cf., e.g., [18–20]). We assume a periodical distribution of the composite microstructure treated as an infinitesimal square tracheidal periodicity cell consisting of an interior part with a void (no material) surrounded by a layer of SiC and an outer layer of carbon. Suppose that each constituent in the microcell consists of an isotropic and homogeneous material. Based on the Hooke law as the constitutive equation, the macroscopic homogenized model is obtained by an asymptotic expansion of the solution of the nonhomogenized elasticity equation with a small scale parameter (see [17]).

We concentrate our efforts on the problem to compute the optimal distribution of our composite material in a given domain. Let $\Omega \subset \mathcal{R}^2$ be a suitable chosen domain that allows to introduce a surface traction t applied to $\Gamma_T \subset \partial\Omega$. On the remaining portion Γ_D of the boundary the displacements are specified.

As an objective function, we consider the mean compliance of the structure defined as follows:

$$\min_{u,\sigma} f(u,\sigma) = \min_{u,\sigma} \left(\int_{\Omega} q \cdot u \, dx + \int_{\Gamma_T} t \cdot u \, ds \right), \tag{10}$$

where q is the external body force applied to Ω. The displacement vector $u = (u_1, u_2)^T$ represents the state variables, and the vector $\sigma = (\sigma_1, \sigma_2)^T$ stands for the design parameters (lengths of the layers in the microstructure). We assume a design composite with a square hole located at the center of the unit microcell and suppose that the lengths and widths of the corresponding material regions are equal. We have denoted by σ_1 the length of the carbon layer and by σ_2 the length of the silicon carbide layer in the microstructure. This assumption is advantageous because it uses less variables in the optimization problem for the material distribution.

The minimization problem (10) is subjected to the following equality and inequality constraints:

$$\sum_{i,j,k,l=1}^{2} \int_{\Omega} E^H_{ijkl} \frac{\partial u_k}{\partial x_l} \frac{\partial \phi_i}{\partial x_j}\, dx = \int_{\Omega} q \cdot \phi \, dx + \int_{\Gamma_T} t \cdot \phi \, ds \qquad \forall \phi \in V_0, \tag{11}$$

$$g(\sigma) = \sum_{i=1}^{2} \sigma_i = C, \qquad \sigma_{\min}\, e \leqslant \sigma \leqslant \sigma_{\max}\, e, \tag{12}$$

for $\sigma_{\min} = 0,\ \sigma_{\max} = 0.5$ and $e = (1,1)^T$. Note that $\sigma_i = 0,\ i = 1,2$, corresponds to a complete void, $\sigma_1 + \sigma_2 = 0.5$ corresponds to a complete solid material, and $0 < \sigma_1, \sigma_2 < 0.5$ and $0 < \sigma_1 + \sigma_2 < 0.5$ to the porous composite with a void at a microlevel. The parameters E^H_{ijkl} in (11) are called *effective material coefficients* for the composite. These quantities can be obtained analytically or numerically (for more complex structures) through a suitable microstructure (see, e.g., [17]). The space V_0 is defined as $V_0 = \{v \in H^1(\Omega) | v = 0 \text{ on } \Gamma_D\}$.

The discretized nonlinear constrained optimization problem has the following form:

$$\min_{u,\sigma} f(u, \sigma), \tag{13}$$

subject to the following constraints:

$$A(\sigma)\, u - b = 0, \quad g(\sigma) - \mathrm{C} = 0, \quad \sigma - \sigma_{\min} e \geqslant 0, \quad \sigma_{\max} e - \sigma \geqslant 0, \tag{14}$$

where $f(u, \sigma)$ is defined by (10), $A(\sigma)$ is the stiffness matrix corresponding to the homogenized equilibrium equation (11), and b is the discrete load vector.

3. Primal-dual Newton interior-point method

3.1. General nonlinear optimization problem

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

$$\min_{x \in \mathcal{R}^n} f(x), \tag{15}$$

subject to

$$h(x) = 0, \quad g(x) \geqslant 0, \tag{16}$$

where $f : \mathcal{R}^n \to \mathcal{R},\ h : \mathcal{R}^n \to \mathcal{R}^m, m < n$, and $g : \mathcal{R}^n \to \mathcal{R}^l$ are assumed to be twice Lipschitz continuously differentiable. Note that $h(x) = 0$ and $g(x) \geqslant 0$ represent vectors $h_i(x) = 0, i = 1, \ldots, m$ and $g_i(x) \geqslant 0,\ i = 1, \ldots, l$, i.e., the equality and inequality constraints have to be understood componentwise.

The *Lagrangian function* associated with (15)–(16) is defined by

$$\mathcal{L}(x, y, z) = f(x) + y^T h(x) - z^T g(x), \tag{17}$$

where y is an $m-$vector and z is an $l-$vector, the Lagrange multipliers for the equality and inequality constraints.

The first-order Karush-Kuhn-Tucker (KKT) necessary conditions for optimality of (15)–(16) read

$$\nabla_x \mathcal{L}(x,y,z) = 0,\ h(x) = 0,\ g(x) \geqslant 0,\ Z\,g(x) = 0,\ z \geqslant 0,$$

where

$$\nabla_x \mathcal{L}(x,y,z) = \nabla f(x) + \sum_{i=1}^{m} y_i \nabla h_i(x) - \sum_{i=1}^{l} z_i \nabla g_i(x) \tag{18}$$

is the gradient of the Lagrangian function and Z is the diagonal matrix with diagonal z. We also consider the Hessian of the Lagrangian with respect to x defined by

$$\nabla_x^2 \mathcal{L}(x,y,z) = \nabla^2 f(x) + \sum_{i=1}^{m} y_i \nabla^2 h_i(x) - \sum_{i=1}^{l} z_i \nabla^2 g_i(x), \tag{19}$$

where $\nabla^2 f(x)$, $\nabla^2 h_i(x)$, $1 \leqslant i \leqslant m$, $\nabla^2 g_i(x)$, $1 \leqslant i \leqslant l$ stand for the Hessinas of $f(x), h_i(x)$, and $g_i(x)$, respectively. Denote by

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

the set of all indices for which the inequality constraints are equal to zero at x. We are interested in finding local minimizers of our optimization problem (15)–(16). Assume that at least one such point x^* exists satisfying the conditions:

(C1) **Feasibility.** $h(x^*) = 0$ and $g(x^*) \geqslant 0$.

(C2) **Regularity.** The set $\{\nabla h_1(x^*),\ldots,\nabla h_m(x^*)\} \cup \{\nabla g_i(x^*),\ i \in \mathcal{A}(x^*)\}$ of gradients of equality and active inequality constraints is linearly independent.

(C3) **Smoothness.** The Hessian matrices $\nabla^2 f(x)$, $\nabla^2 h_i(x)$, $1 \leqslant i \leqslant m$, and $\nabla^2 g_i(x)$, $1 \leqslant i \leqslant l$, exist and are locally Lipschitz continuous at x^*.

(C4) **Second-order sufficiency condition.** $\eta^T \nabla_x^2 \mathcal{L}(x^*)\eta > 0$ for all vectors $\eta \neq 0$ satisfying $\nabla h_i(x^*)^T \eta = 0$, $1 \leqslant i \leqslant m$, and $\nabla g_i(x^*)^T \eta = 0$, $i \in \mathcal{A}(x^*)$.

(C5) **Strict complementarity.** $z_i^* > 0$ if $g_i(x^*) = 0$, $1 \leqslant i \leqslant l$.

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 $g(x) \geqslant 0$ can be replaced by $g(x) - s = 0$, $s \geqslant 0$ 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., [4,5,8,26]. 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. Recently, this rather popular approach was applied for solving constrained optimal control problems (see, e.g., [2]). The third approach was used in our practical implementations and we explain it in detail in the next subsection.

3.2. Logarithmic barrier interior-point method

The logarithmic barrier function method was first introduced by Frisch [10] and later on popularized by Fiacco and McCormick [9] in the late sixties of the last century. The basic idea of this method is to replace the optimization problem (15)–(16) with the following equality constrained optimization problem:

$$\min_{x\in\mathcal{R}^n} \beta^{(\rho)}(x), \tag{20}$$

subject to

$$h(x) = 0, \tag{21}$$

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

$$\beta^{(\rho)}(x) = f(x) - \rho \sum_{i=1}^{l} \log\, g_i(x) \tag{22}$$

is often referred to as a *barrier function*. To insure the existence of logarithmic terms in (22), we implicitly require $g_i(x) > 0,\ 1 \leqslant i \leqslant l$. In such a way, we get a family of subproblems depending on ρ for which it is well-known that under the conditions (C1)–(C5) from Section 3.1 the solution of (20)–(21) reduces to the solution of the original problem (15)–(16) as ρ decreases to zero (cf. [9]). This method obviously is an interior-point method, because 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 $x^{(\rho)}$ such that $g(x^{(\rho)}) > 0$. The solution points $x^{(\rho)}$ parametrized by ρ define the so-called *central path* or also called the *barrier trajectory*.

The gradient of (22) is given by

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

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

$$\nabla^2\beta^{(\rho)}(x) = \nabla^2 f(x) - \sum_{i=1}^{l} \frac{\rho}{g_i(x)} \nabla^2 g_i(x) + \sum_{i=1}^{l} \frac{\rho}{g_i^2(x)} \nabla g_i(x)(\nabla g_i(x))^T. \tag{23}$$

The Lagrangian function associated with (20)–(21) is

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

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

$$\nabla_x \mathcal{L}^{(\rho)}(x, y) = \nabla\beta^{(\rho)}(x) + \nabla h(x) y = \nabla f(x) - \sum_{i=1}^{l} \frac{\rho}{g_i(x)} \nabla g_i(x) + \sum_{i=1}^{m} y_i \nabla h_i(x). \tag{24}$$

The logarithmic barrier function method consists now of generating a sequence of iterative solutions $\{x\} = \{x^{(\rho)}\}$, local minimizers of the equality constrained subproblems, with $\rho > 0$

decreasing at each iteration. Taking into account the first-order optimality conditions and especially $\nabla_x \mathcal{L}^{(\rho)}(x,y) = 0$, we see that the convergence of $\{x^{(\rho)}\}$ to an optimal solution x^* requires that

$$\lim_{\rho \to 0} y_i^{(\rho)} = y_i^*,\ 1 \leqslant i \leqslant m \qquad \text{and} \qquad \lim_{\rho \to 0} \frac{\rho}{g_i(x^{(\rho)})} = z_i^*,\ 1 \leqslant i \leqslant l, \tag{25}$$

where $\{y_i^*\}$ and $\{z_i^*\}$ are the Lagrange multipliers associated with the equality and inequality constraints $g_i(x^{(\rho)}) > 0$, respectively. From $g_i(x^{(\rho)}) \to 0$ and the second relation in (25) we get $\rho / g_i^2(x^{(\rho)}) \to \infty$ and hence, the Hessian of the logarithmic barrier function (23) would become arbitrarily large. Comparing now relations (18) and (24), we see that $\rho / g_i(x^{(\rho)})$ serves as a Lagrange multiplier for the inequality constraints. Thus, we can introduce an auxiliary variable $z_i = z_i^{(\rho)} = \rho / g_i(x^{(\rho)}),\ 1 \leqslant i \leqslant l$ which can also be written in the form $z_i^{(\rho)} g_i(x^{(\rho)}) = \rho$. 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 now formulate the perturbed KKT conditions for the logarithmic barrier function problem (20)–(21), namely,

$$\nabla f(x) + \nabla h(x) y - \nabla g(x) z = 0,\ h(x) = 0,\ Zg(x) = \rho,\ g(x) > 0. \tag{26}$$

In matrix-vector notations, (26) results in the following nonlinear equation with $n + m + l$ components:

$$F^{(\rho)}(x,y,z) = 0 \quad \text{with} \quad F^{(\rho)}(x,y,z) = \begin{pmatrix} t + J_{\text{eq}}^T y - J_{\text{in}}^T z \\ h \\ Gz - \rho\, e \end{pmatrix}, \tag{27}$$

where $F^{(\rho)} = \nabla \mathcal{L}^{(\rho)}$ is the gradient of the Lagrangian function with respect to x, y, and z; $t = \nabla f(x)$ is the gradient of the objective function, J_{eq} is the Jacobian $m \times n$ matrix of the equality constraints $h(x) = 0$, and J_{in} is the Jacobian $l \times n$ matrix of the inequality constraints $g(x) \geqslant 0$. In the last equation of (27) we have denoted $G = \text{diag}(g_i),\ g_i > 0,\ 1 \leqslant i \leqslant l$ and $e = (1, 1, \ldots, 1)^T$. Note that at each iteration we have three sets of unknowns: the primal variable x, the dual variable y, and the perturbed complementarity variable z which we consider independently.

Denote now the unknown solution by $\Phi = (x, y, z)$ and apply the Newton method to the nonlinear system (27) to find the increments $\Delta\Phi = (\Delta x, \Delta y, \Delta z)$, namely,

$$K \Delta\Phi = -F^{(\rho)}(\Phi), \tag{28}$$

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

$$K = \begin{pmatrix} H & J_{\text{eq}}^T & -J_{\text{in}}^T \\ J_{\text{eq}} & 0 & 0 \\ Z J_{\text{in}} & 0 & G \end{pmatrix}, \tag{29}$$

where the Hessian of the Lagrangian function $H = \nabla_x^2 \mathcal{L}$ is given by (19). 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 [11, 12, 25] for more details). In the

case of convex optimization (i.e., convex objective function $f(x)$, linear equality constraints $h(x)$, and concave inequality constraints $g(x)$), the Hessian matrix H is positive semidefinite. The properties of the Hessian matrix for the inequality constrained optimization problem with a logarithmic barrier function method are discussed in [24, 25].

One possible way for solving (28) is to symmetrize K taking into account the fact that Z and G are diagonal matrices. This method is proposed in [12] and results in the following symmetric matrix:

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

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 [12] 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 (28) which we use in our practical applications is to eliminate the (1,3) block of (29), i.e., due to $g(x) > 0$, we eliminate Δz from the third equation of (28)

$$\Delta z = -z + G^{-1}(\rho\, e - Z J_{\text{in}} \Delta x) \tag{30}$$

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

$$\begin{pmatrix} \tilde{H} & J_{\text{eq}}^T \\ J_{\text{eq}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = - \begin{pmatrix} t + J_{\text{eq}}^T y - \rho J_{\text{in}}^T G^{-1} e \\ h \end{pmatrix}, \tag{31}$$

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 (31) is called a *condensed* primal-dual system. A detailed analysis of the properties of the condensed primal-dual matrix can be found in [25] 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 (31) and finding the primal-dual steps $(\Delta x, \Delta y)$ are proposed in the literature (cg., e.g., [8, 13, 24]). Note that one needs a reliable and efficient solver of (31), since the condensed primal-dual system is solved at every iteration of the optimization loop. In practice, we apply transforming iterations (see [28]) to find the increments. This method will be explained in more detail in Section 4.

After finding the solution of (31), the algorithm proceeds iteratively from an initial point $(x^{(0)}, y^{(0)}, z^{(0)})$ through a sequence of points determined from the search directions described by (30) and (31) as follows:

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

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.

3.3. 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., [8,13,26]). 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., [4,5,7]. Typical of this method is to find a step $d^{(k)}$ which is restricted to a set, called the *trust region*. This set is practically obtained by limiting $|d^{(k)}| \leqslant r^{(k)}$, 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 $d^{(k)}$, update the solution $\Phi^{(k+1)} = \Phi^{(k)} + d^{(k)}$ 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 $d^{(k)}$.

We apply the second variant of the Newton method, the line-search approach. Once the solution $\Delta\Phi^{(k)}$ of (28) has been determined, we find a steplength $\alpha^{(k)} > 0$ such that $\Phi^{(k+1)} = \Phi^{(k)} + \alpha^{(k)}\Delta\Phi^{(k)}$ measuring a progress in minimization at each iteration and reducing the merit function in the sense $M(\Phi^{(k+1)}) < M(\Phi^{(k)})$. 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 $\Phi^{(k)}$ as the solution.

S2. Compute a search direction $\Delta\Phi^{(k)}$ solving (28).

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

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

A standard convergence monitor in nonlinear programming is to choose the Euclidean norm $\|F^{(\rho)}(x, y, z)\|$ of the residual produced by the KKT conditions (27) 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 nonminimizing point. The KKT conditions are necessary optimality conditions and, hence, the optimization problem (15)–(16) and the nonlinear problem (27) are not equivalenet, i.e., the Newton method may find solutions of (27) which do not minimize the objective function $f(x)$. Therefore, in order to find simultaneously solutions of both problems, a better approach is to rely on the hierarchy of two merit functions (cf., e.g., [13,16]). 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., [5,8,11,26]). In particular, our *primary merit function* is related to those suggested in [13] and is chosen as a modified augmented Lagrangian incorporating the logarithmic

barrier function (22) as follows:

$$M = M(x, y, \rho, \rho_A) = f(x) - \rho \sum_{i=1}^{l} \log\, g_i(x) + y^T h(x) + \frac{1}{2}\rho_A\, h(x)^T h(x), \tag{32}$$

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 (20)–(21). Note that a descent is sought only with respect to x taking into account the original optimization problem. A standard way to achieve $M(x + \alpha\Delta x, y, \rho, \rho_A) < M(x, y, \rho, \rho_A)$ is to require that Δx is a descent direction, i.e., $\Delta x^T \nabla_x M < 0$, where $\nabla_x M$ is the gradient of the primary merit function with respect to the primal variable x. In particular, we have

$$\begin{aligned} \Delta x^T \nabla_x M &= \Delta x^T (t - \rho J_{\text{in}}^T G^{-1} e + J_{\text{eq}}^T y + \rho_A J_{\text{eq}}^T h) \\ &= \Delta x^T (t - \rho J_{\text{in}}^T G^{-1} e) - h^T y - \rho_A h^T h, \end{aligned} \tag{33}$$

due to $J_{\text{eq}}\Delta x = -h$ from the second equation of (31). Hence, $\Delta x^T \nabla_x M < 0$ can be satisfied if the augmented Lagrangian parameter ρ_A is sufficiently large, namely,

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

Hence, ρ_A can be changed within the optimization loop if Δx is not a descent direction. In our algorithm, we choose

$$\rho_A = \min\left(\frac{5}{h^T h}(\Delta x^T (t - \rho J_{\text{in}}^T G^{-1} e) - h^T y), 100\right) \tag{34}$$

in the case of $\Delta x^T \nabla_x M \geqslant 0$ and continue the loop (see [13, 16] for details).

For the *secondary merit function* we choose the l_2- norm of the residual with respect to the perturbed KKT-conditions (27). We apply the Newton method and choose the steplengths to strictly satisfy the inequality constraints $g(x) > 0$ and the complementarity constraints $z > 0$. Hence, the first requirement for the line-search approach is to ensure a strict feasibility. Let $\hat{\alpha}$ and $\hat{\gamma}$ be separate steplengths for x and z defined as follows:

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

Since we maintain interior (i.e., strictly feasible) iterates we usually 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 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., [8, 13, 16]). We rely on a *watchdog strategy* (see [6]) 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 (34).

4. Solving the condensed primal-dual system

The discretized constrained minimization problems (8)–(9) and (13)–(14) are solved by the primal-dual interior-point method described in Section 3. We consider the diagonal matrices $D_1 = \mathrm{diag}(\sigma_i - \sigma_{\min})$ and $D_2 = \mathrm{diag}(\sigma_{\max} - \sigma_i)$ and introduce $z = \rho D_1^{-1} e \geqslant 0$ and $w = \rho D_2^{-1} e \geqslant 0$ serving as a perturbed complementarity. Note that for the problem described in Section 2.1 we have $1 \leqslant i \leqslant N$, where N is the number of finite elements in the discretized domain and $e = (e_1, \ldots, e_N)^T$, $e_i = 1$. For the problem in Section 2.2 we have, respectively, $1 \leqslant i \leqslant 2$ and $e = (1,1)^T$. The primal-dual Newton-type interior-point method is applied to three sets of variables: primal feasibility (u, σ), dual feasibility (λ, η), and perturbed complementarity related to (z, w).

Denote the Lagrangian function of (8)–(9), respectively (13)–(14), by

$$\begin{aligned}\mathcal{L}(u, \sigma; \lambda, \eta; z, w) = f(u, \sigma) + \lambda^T (A(\sigma)\, u - b) + \eta\, (g(\sigma) - C) \\ - z^T (\sigma - \sigma_{\min}\, e) - w^T (\sigma_{\max} e - \sigma).\end{aligned} \tag{35}$$

The Newton method applied to the KKT conditions of (35) results in

$$\begin{pmatrix} 0 & \mathcal{L}_{u\sigma} & \mathcal{L}_{u\lambda} & 0 & 0 & 0 \\ \mathcal{L}_{\sigma u} & \mathcal{L}_{\sigma\sigma} & \mathcal{L}_{\sigma\lambda} & \mathcal{L}_{\sigma\eta} & -I & I \\ \mathcal{L}_{\lambda u} & \mathcal{L}_{\lambda\sigma} & 0 & 0 & 0 & 0 \\ 0 & \mathcal{L}_{\eta\sigma} & 0 & 0 & 0 & 0 \\ 0 & Z & 0 & 0 & D_1 & 0 \\ 0 & -W & 0 & 0 & 0 & D_2 \end{pmatrix} \begin{pmatrix} \triangle u \\ \triangle \sigma \\ \triangle \lambda \\ \triangle \eta \\ \triangle z \\ \triangle w \end{pmatrix} = - \begin{pmatrix} \nabla_u \mathcal{L} \\ \nabla_\sigma \mathcal{L} \\ \nabla_\lambda \mathcal{L} \\ \nabla_\eta \mathcal{L} \\ \nabla_z \mathcal{L} \\ \nabla_w \mathcal{L} \end{pmatrix}, \tag{36}$$

where I stands for the identity matrix, $Z = \mathrm{diag}(z_i)$ and $W = \mathrm{diag}(w_i)$ are diagonal matrices. Following Section 3, we eliminate the increments for z and w from the 5th and 6th rows of (36), namely,

$$\triangle z = D_1^{-1}(-\nabla_z \mathcal{L} - Z \triangle \sigma), \qquad \triangle w = D_2^{-1}(-\nabla_w \mathcal{L} + W \triangle \sigma) \tag{37}$$

and substitute (37) in the second row of (36). We get the linear system $\tilde{K} \triangle \psi = -\tilde{\xi}$ for the increments of $\psi = (u, \sigma, \lambda, \eta)$, denoted by $\triangle \psi = (\triangle u, \triangle \sigma, \triangle \lambda, \triangle \eta)$, where $\tilde{K}$ is the matrix and $(-\tilde{\xi})$ is the right-hand side of the following condensed primal-dual system:

$$\begin{pmatrix} 0 & \mathcal{L}_{u\sigma} & \mathcal{L}_{u\lambda} & 0 \\ \mathcal{L}_{\sigma u} & \tilde{\mathcal{L}}_{\sigma\sigma} & \mathcal{L}_{\sigma\lambda} & \mathcal{L}_{\sigma\eta} \\ \mathcal{L}_{\lambda u} & \mathcal{L}_{\lambda\sigma} & 0 & 0 \\ 0 & \mathcal{L}_{\eta\sigma} & 0 & 0 \end{pmatrix} \begin{pmatrix} \triangle u \\ \triangle \sigma \\ \triangle \lambda \\ \triangle \eta \end{pmatrix} = - \begin{pmatrix} \nabla_u \mathcal{L} \\ \tilde{\nabla}_\sigma \mathcal{L} \\ \nabla_\lambda \mathcal{L} \\ \nabla_\eta \mathcal{L} \end{pmatrix}. \tag{38}$$

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

$$\tilde{\mathcal{L}}_{\sigma\sigma} = \mathcal{L}_{\sigma\sigma} + D_1^{-1}\, Z + D_2^{-1}\, W, \qquad \tilde{\nabla}_\sigma \mathcal{L} = \nabla_\sigma \mathcal{L} + D_1^{-1}\, \nabla_z \mathcal{L} - D_2^{-1}\, \nabla_w \mathcal{L}.$$

The direct methods for the solution of (38) can be divided into two classes: *range space methods* and *null space methods* (cf., e.g., [14]). These approaches essentially differ in the grouping of the matrix into a 2×2-block structure. The decomposition of the condensed primal-dual system (38) 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 third rows and columns. The resulting matrix can be written according to

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

where the first diagonal block

$$A_{11} = \begin{pmatrix} 0 & \mathcal{L}_{u\lambda} \\ \mathcal{L}_{\lambda u} & 0 \end{pmatrix} \tag{39}$$

is now an indefinite but nonsingular matrix. We remind that $\mathcal{L}_{\lambda u} = A(\sigma)$ is exactly the stiffness matrix corresponding to the state equation: either the electric potential equation (3) or the equilibrium equation (11). 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 $\tilde{K}$:

$$K^L \tilde{K} K^R = M_1 - M_2 \tag{40}$$

with the left and right factors given below and reasonable matrices M_1 and $M_2 \sim 0$. To solve the system of the form $\tilde{K}\triangle\psi = -\tilde{\xi}$, starting with an initial guess for $\triangle\psi = (\triangle u, \triangle\lambda, \triangle\sigma, \triangle\eta)^T$, the transforming iteration proposed in [28] is described by

$$\triangle\psi^{(\nu+1)} = \triangle\psi^{(\nu)} + K^R M_1^{-1} K^L(-\tilde{\xi} - \tilde{K}\triangle\psi^{(\nu)}), \tag{41}$$

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

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

The line-search approach and the choice of the steplengths parameters α_j are discussed in Section 3.3.

Using an appropriate preconditioner for the stiffness matrix, we approximate the first diagonal block (39) as follows:

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

Usually, the left and right transformations are of the form

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

In this case, the regular splitting (40) becomes $KK^R = M_1 - M_2$, where

$$M_1 = \begin{pmatrix} 0 & \mathcal{L}_{u\lambda} & 0 & 0 \\ \mathcal{L}_{\lambda u} & 0 & 0 & 0 \\ \mathcal{L}_{\sigma u} & \mathcal{L}_{\sigma\lambda} & \tilde{S} & \mathcal{L}_{\sigma\eta} \\ 0 & 0 & \mathcal{L}_{\eta\sigma} & 0 \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ R & Q \end{pmatrix} \tag{42}$$

and

$$M_2 = \begin{pmatrix} 0 & 0 & \mathcal{L}_{u\sigma} - \mathcal{L}_{u\lambda}\tilde{\mathcal{L}}_{u\lambda}^{-1}\mathcal{L}_{u\sigma} & 0 \\ 0 & 0 & \mathcal{L}_{\lambda\sigma} - \mathcal{L}_{\lambda u}\tilde{\mathcal{L}}_{\lambda u}^{-1}\mathcal{L}_{\lambda\sigma} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{43}$$

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 $\mathcal{L}_{u\lambda}$. The second diagonal block Q in (42) is symmetric and indefinite given by

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

We denote the defect in (41) by $d = -\tilde{\xi} - \tilde{K}\triangle\psi^{(\nu)}$ and compute the corresponding entries:

$$d_u = -\nabla_u\mathcal{L} - \mathcal{L}_{u\lambda}\triangle\lambda - \mathcal{L}_{u\sigma}\triangle\sigma, \qquad d_\lambda = -\nabla_\lambda\mathcal{L} - \mathcal{L}_{\lambda u}\triangle u - \mathcal{L}_{\lambda\sigma}\triangle\sigma,$$
$$d_\sigma = -\tilde{\nabla}_\sigma\mathcal{L} - \mathcal{L}_{\sigma u}\triangle u - \mathcal{L}_{\sigma\lambda}\triangle\lambda - \tilde{\mathcal{L}}_{\sigma\sigma}\triangle\sigma - \mathcal{L}_{\sigma\eta}\triangle\eta, \qquad d_\eta = -\nabla_\eta\mathcal{L} - \mathcal{L}_{\eta\sigma}\triangle\sigma.$$

Taking into account (41), one needs to compute $\delta = M_1^{-1}d$, i.e., $M_1\delta = d$. Consequently, we find $\delta_\lambda = \tilde{\mathcal{L}}_{u\lambda}^{-1}d_u$ and $\delta_u = \tilde{\mathcal{L}}_{\lambda u}^{-1}d_\lambda$. To compute the remaining components of δ, we have to solve systems with an indefinite matrix Q of the form

$$\begin{pmatrix} \tilde{S} & \mathcal{L}_{\sigma\eta} \\ \mathcal{L}_{\eta\sigma} & 0 \end{pmatrix} \begin{pmatrix} \delta_\sigma \\ \delta_\eta \end{pmatrix} = \begin{pmatrix} d_\sigma - L_{\sigma u}\delta_u - L_{\sigma\lambda}\delta_\lambda \\ d_\eta \end{pmatrix}.$$

Iterative procedures such as MINRES or Bi-CGSTAB (see [27]) with appropriate stopping criteria can be applied in this case. Compute $K^R\delta$ and find the increments from (41) as follows:

$$\triangle u^{(\text{new})} = \triangle u^{(\text{old})} + \delta_u - \tilde{\mathcal{L}}_{\lambda u}^{-1}\mathcal{L}_{\lambda\sigma}\delta_\sigma, \quad \triangle\sigma^{(\text{new})} = \triangle\sigma^{(\text{old})} + \delta_\sigma,$$
$$\triangle\lambda^{(\text{new})} = \triangle\lambda^{(\text{old})} + \delta_\lambda - \tilde{\mathcal{L}}_{u\lambda}^{-1}\mathcal{L}_{u\sigma}\delta_\sigma, \quad \triangle\eta^{(\text{new})} = \triangle\eta^{(\text{old})} + \delta_\eta.$$

We apply the above algorithm (with a fixed number of iterations) to find the increments of the primal and dual variables $\triangle u, \triangle\sigma, \triangle\lambda, \triangle\eta$ and then use (37) to determine the global search direction $\triangle\Phi$.

5. Numerical experiments

In this section, we present the results of some numerical experiments with solving the optimization problems associated with our two applications. First, we solve the problem (8)–(9). Note that instead of the conductivity in (3), we consider some modification of the form

$$\operatorname{div}(h(\sigma)\operatorname{grad}\varphi) = 0 \text{ in } \Omega, \quad \mathrm{n}\cdot h(\sigma)\operatorname{grad}\varphi = \begin{cases} \mathrm{I}_\nu & \text{on } \Gamma_\nu \subset \partial\Omega, \\ 0 & \text{elsewhere,} \end{cases} \tag{44}$$

where

$$h(\sigma) = \frac{\sigma - \sigma_{\min} + \varepsilon}{\sigma_{\max} - \sigma_{\min}}, \qquad 0 < \varepsilon \ll 1. \tag{45}$$

We run our optimization algorithm (see [16] for details) with the initial lower and upper bounds for the conductivity $\sigma_{\min} = 0.01$ and $\sigma_{\max} = 1$, respectively, and take $\varepsilon = 0.01$. The

numerical results treat a two-dimensional isotropic material occupying a rectangular domain decomposed into $N = N_x\, N_y$ quadrilateral finite elements, where N_x and N_y are the number of grid points on the coordinate axes Ox and Oy. Note that the conductivity is a vector with N components and we find the corresponding entries computing the conductivity at the center points of the finite elements. As an initial distribution we take $\sigma_i = 0.45,\ i = 1, \ldots, N$.

Table 1. Convergence results for Maxwell equations

N_c	N_x	N_y	iter	ρ	M	$\|F^{(\rho)}\|_2$	$\|v\|_2$
2	10	10	15	4.21e-17	3.57	6.94e-6	1.0e-9
2	30	30	21	2.64e-18	4.25	1.83e-4	1.0e-9
2	30	40	19	1.97e-17	4.47	3.49e-5	1.0e-9
2	40	40	28	7.56e-17	4.73	3.40e-4	1.0e-9
2	60	80	38	7.76e-17	4.79	5.62e-5	1.0e-9
3	30	40	33	7.03e-18	3.41	7.62e-4	1.0e-10
3	50	50	30	6.44e-19	3.78	2.40e-4	1.0e-11
3	60	70	37	3.48e-17	3.89	6.84e-4	1.0e-10
3	80	80	43	5.21e-18	3.95	5.32e-4	1.0e-10
4	30	30	19	3.21e-17	24.75	5.12e-5	1.0e-10
4	50	60	21	7.34e-18	27.29	8.12e-5	1.0e-11
5	30	40	19	1.56e-17	68.39	6.72e-4	1.0e-9
5	70	70	22	7.48e-18	79.14	8.45e-4	1.0e-10
6	50	50	20	9.85e-17	81.99	1.27e-3	1.0e-9
6	60	80	27	8.12e-17	84.03	5.17e-4	1.0e-9

In Table 1, we present the values of N_c, the number of contacts (assuming that the sum of all currents on the boundary is equal to zero), N_x and N_y, the number of iterations iter to achieve convergence, the last value of the barrier parameter ρ, the final value of the primary merit function M in view of (32), the norm of the residual $\|F^{(\rho)}\|$ as a secondary merit function, and the l_2-norm $v = (z, w)$ related to the complementarity conditions at the last iteration.

Fig. 1 displays the material distribution for a mesh 50×50 with 3 and 6 contacts, respectively. On the Oz-direction the values of the conductivity are visualized. We see that the high profiles correspond to σ close to $\sigma_{\max} = 1$ and the low values of σ correspond to the part of the domain with a void (no material).

We are now concerned with the solution of problem (13)–(14). In Table 2 we report some numerical results from running the optimization code varying the constant C with respect to (12). Our purpose is to find the optimal lengths of the layers in the composite material and to show the convergence behavior of the optimization algorithm. In this case, the domain is generated by MATLAB and is chosen to be a circle which corresponds naturally to the 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 $\sigma_1^{(0)}$ and $\sigma_2^{(0)}$. As before, we report the number of iterations iter to get convergence, the optimal lengths σ_1 and σ_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 $v = (z, w)$ at the last iteration. We see from the experiments that the optimal length σ_1 of the carbon layer in all the runs is very close to zero, i.e., the

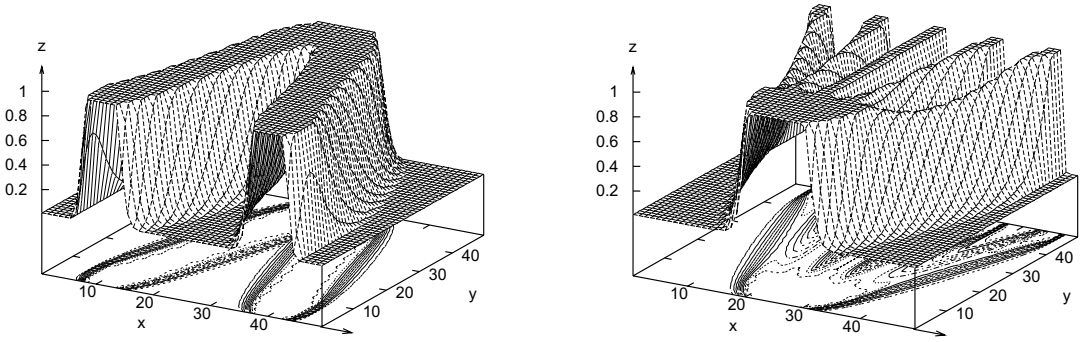

Figure 1. Material distribution: 50×50 mesh; a) 3 contacts; b) 6 contacts

Table 2. Convergence results for biomorphic microcellular SiC ceramics

$\sigma_1^{(0)}$	$\sigma_2^{(0)}$	C	iter	σ_1	σ_2	ρ	M	$\|F^{(\rho)}\|_2$	$\|v\|_2$
0.05	0.05	0.3	11	3.66e-12	0.3	1.35e-17	1.24	9.63e-6	1.0e-10
0.1	0.1	0.3	11	5.51e-14	0.3	3.09e-21	1.24	1.03e-6	1.0e-12
0.1	0.1	0.4	12	1.66e-16	0.4	1.29e-26	0.85	8.63e-9	1.0e-14
0.2	0.2	0.1	16	5.55e-17	0.1	2.23e-25	7.73	2.23e-8	1.0e-13
0.2	0.2	0.2	13	1.08e-16	0.2	5.35e-26	2.34	1.54e-8	1.0e-14
0.2	0.2	0.3	11	2.58e-16	0.3	6.76e-26	1.24	1.79e-8	1.0e-14
0.24	0.24	0.15	11	5.41e-15	0.15	4.14e-12	3.81	4.99e-7	1.0e-12
0.3	0.1	0.4	11	1.35e-12	0.4	8.50-19	0.85	5.07e-6	1.0e-10
0.4	0.05	0.1	17	9.83e-15	0.1	6.92e-21	7.73	9.49e-7	1.0e-11

solid part of the body is entirely occupied by a silicon carbide layer due to the higher stiffness of this material.

Our stopping criterion for the optimization algorithm in both applications is

$$\|F^{(\rho)}\| < \mathsf{tol} \quad or \quad \rho < \mathsf{tol}^2 \quad or \quad \mathsf{iter} > \mathsf{itmax} \quad or \quad \mathsf{lsteps} > \mathsf{lsmax},$$

where $\mathsf{tol} = 10^{-8}$, the maximum number of iterations in the main optimization loop is $\mathsf{itmax} = 100$, and the specified limit on the number of iterations in the line search loop is $\mathsf{lsmax} = 15$. In both cases, we have chosen the initial value for the barrier parameter $\rho = 1$ and the initial value for the augmented Lagrangian parameter $\rho_A = 10$.

Conclusions

Summarizing, we have considered two structural optimization problems in materials science. The first one deals with the topology optimization in conductive electromagnetic media described by the Maxwell equations where the design optimization is to minimize the energy dissipation caused by eddy currents. The second one is concerned with the shape optimization of the biomorphic microcellular SiC ceramics, recently produced from natural wood, based on homogenization modeling. The design objective for the latter problem is to determine the microstructural geometric features in order to get an optimal performance of

the composite materials under consideration. In both cases, the objective function is subjected to equality and inequality constraints on the state and design parameters typically leading to a constrained nonlinear minimization problem exhibiting a variety of local optima and saddle points. We have used the primal-dual Newton interior-point method to find the increments. Convergence is achieved by using the watchdog technique based on the hierarchy of appropriately chosen merit functions. For both structural optimization problems, we have included numerical experiments to demonstrate the respective material distributions in two-dimensional domains.

Acknowledgements

This work has been partially supported by the German National Science Foundation (DFG) under Grant No. HO877/5-2. The second author has also been supported in part by the Bulgarian Ministry for Education and Science under Grant I1001/2000.

References

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

[2] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch, *A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems*, SIAM J. Optim., **11** (2000), pp. 495–521.

[3] O. Biro and K. Preis, *Various FEM formulations for the calculation of transient 3d eddy currents in nonlinear media*, IEEE Transactions on Magnetics, **31** (1995), pp. 1307–1312.

[4] R. H. Byrd, J. C. Gilbert, and J. Nocedal, *A trust region method based on interior point techniques for nonlinear programming*, Math. Program., Ser. A, **9** (2000), pp. 149–185.

[5] R. H. Byrd, M. E. Hribar, and J. Nocedal, *An interior point algorithm for large scale nonlinear programming*, SIAM J. Optim., **9** (1999), No. 4, pp. 877–900.

[6] 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. Program. Study, **16** (1982), pp. 1–17.

[7] A. R. Conn, N. I. M. Gould, D. Orban, P. L. Toint, *A primal-dual trust-region algorithm for non-convex nonlinear programming*, Math. Program., Ser. B, **87** (2000), pp. 215–249.

[8] 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. Optimization Theory Appl., **89** (1996), pp. 507–541.

[9] 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.

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

[11] A. Forsgen and P. E. Gill, *Primal-dual interior methods for nonconvex nonlinear programming*, SIAM J. Optim., **8** (1998), pp. 1132–1152.

[12] A. Forsgen, 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** (1996), pp. 187–211.

[13] D. M. Gay, M. L. Overton, and M. H. Wright, *A Primal-Dual Interior method for nonconvex nonlinear programming*, Advances in Nonlinear Programming (1998), Edited by Y. Yuan, Kluwer, Dordrecht, Holland, pp. 31–56.

[14] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, London, 1981.

[15] P. Greil, T. Lifka, and A. Kaindl, *Biomorphic cellular silicon carbide ceramics from wood: I. Processing and microstructure, and II. Mechanical properties*, J. Europ. Ceramic Soc., **18** (1998), pp. 1961–1973 and 1975–1983.

[16] R. H. W. Hoppe, S. I. Petrova, and V. Schulz, *Primal-dual Newton-type interior-point method for topology optimization*, J. Optimization Theory Appl., **114** (2002), No. 3, pp. 545–571.

[17] R. H. W. Hoppe and S. I. Petrova, *Structural optimization of biomorphic microcellular ceramics by homogenization approach*, Lect. Notes Comput. Sci., Springer, **2179** (2001), pp. 353–360.

[18] A. Bensoussan, J. L. Lions, G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North–Holland, Elsevier Science Publishers, Amsterdam, 1978.

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

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

[21] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual method for linear programming*, Progress in Mathematical Programming, Interior Point and Related Methods (1989), Edited by N. Megiddo, Springer Verlag, New York, New York, pp. 29–47.

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

[23] J. Sokolowski and J.-P. Zolesio, *Introduction to Shape Optimization*, Springer Series in Computational Mathematics, **16** (1992), Springer.

[24] M. H. Wright, *Some properties of the Hessian of the logatithmic barrier function*, Math. Program., **67** (1994), pp. 265–295.

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

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

[27] 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** (1992), pp. 631–644.

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

Received 18 Feb. 2003