1 F1309: Multilevel Solvers for Large Scale Discretized Optimization Problems

Title : Multilevel Solvers for Large Scale Discretized Optimization Problems

Principal investigator : o.Univ.-Prof. Dipl.-Ing. Dr. Ulrich Langer
Institute of Computational Mathematics
Johannes Kepler University Linz
Tel.: +43-70-2468-9168
Fax: +43-70-2468-9148
E-Mail: ulanger@numa.uni-linz.ac.at

Co-investigator : o.Univ.-Prof. Dipl.-Ing. Dr. Heinz W. Engl
Industrial Mathematics Institute
Johannes Kepler University Linz
Tel.: +43-70-2468-9219
Fax: +43-70-2468-8855
E-Mail: engl@indmath.uni-linz.ac.at

Co-investigator : Dipl.-Ing. Dr. Ewald Lindner
Institute of Computational Mathematics
Johannes Kepler University Linz
Tel.: +43-70-2468-9165
Fax: +43-70-2468-9148
E-Mail: lindner@numa.uni-linz.ac.at

Disciplinary affiliation:

1	%	2	%	3	%
1114	40	1140	30	1133	30

(according to the Central Bureau of Statistics classification)

1.1 Abstract

Shape optimization tries to modify the shape of an object in such a way that the resulting object is optimal with respect to a certain objective. This goal got more and more important also in industrial applications. In most cases the industrial design process is to be automatized so that the whole design phase is accelerated. In this subproject we focused on the development, analysis and implementation of numerically efficient algorithms for solving optimal design problems including hierarchical shape optimization, see Subsection 1.3.3, and multilevel topology optimization, see Subsection 1.3.1, as well as their applications to real-life problems. Last but not least we investigated the relevant

geometry and mesh handling, see Subsection 1.3.2. Applying similar methods as in topology optimization resulted in efficient simultaneous (all-at-once) methods for parameter identification problems. These results have been obtained in a very fruitful cooperation with Subproject F1308. We further investigated efficient techniques for solving the resulting KKT systems, in particular we are interested in the use of multigrid methods as preconditioners or as an iterative method itself, see Subsection 1.3.4.

Finally, let us summarize the contributions to the goals formulated in the proposal three years ago in the Table below:

Attained Goals	References
Geometry and Mesh Handling	[50], [55], [51], [34], [35]
Hierarchical Shape Optimization Techniques in 3D	[50], [51], [52], [53], [55], [56]
Adaptive Multilevel 3D Topology Optimization Techniques	[97], [96], [98], [26], [25], [95], [99], [100], [51], [52], [53]
KKT Solver	[92], [87], [99], [86]
Synergies	[56], [55], [100], [25], [26], [76], [77]
Ph.D. Theses	[98]
Software	
| NETGEN	www.sfb013.uni-linz.ac.at/software.html
| FEPP	www.sfb013.uni-linz.ac.at/software.html
| NGSolve	www.sfb013.uni-linz.ac.at/software.html
| Optimization tools	www.sfb013.uni-linz.ac.at/software.html

In summary, we published 7 papers in refereed journals, 7 papers in refereed proceedings and 5 technical reports. In addition 1 Ph.D. Thesis were completed during the third funding period.

1.2 Scientific Background and Current State of Research

Optimal design can be seen in a wider context of inverse problems in which we know the behaviour of a system, usually from physical measurements, and using this knowledge we are looking for the structure of the system and/or for the distribution of sources. A typical inverse problem is computer tomography in medicine. An introduction to this field is given in the textbook by Kirsch [47]. Inverse problems are known to be ill–posed, so they have to be treated by regularization techniques, see Engl, Hanke, and Neubauer [32]. Some connections between optimization and inverse problems are presented by Neittaanmäki, Rudnicki, and Savini [62].

In this subproject we are especially interested in structural optimization where we change the structure of an object, which is interacted in some physical equilibrium, in order to achieve a required behaviour. Here structure means either material properties, topology or shape of the object, boundary or interfaces. Various issues of

structural optimization are covered in Banichuk [3], Bendsøe [11], Cherkaev [27], Kalamkarov [44], Olhoff and Taylor [64], Pedersen [66], Rozvany [78, 79, 80], Save and Prager [81, 82], Xie and Steven [110] and Allaire [1]. Applications in plasticity are given, e.g., by Yuge and Kikuchi [112], in electromagnetism, e.g., by Hoppe, Petrova, and Schulz [41], and, for instance, in ergonomics by Rasmussen et al. [74]. An optimal design of microstructures is presented by Jacobsen, Olhoff, and Rønholt [43].

If we are interested in finding topology of a structure – it is usually the question where to put holes and where to put material – we speak about topology optimization. Some of the relevant literature is Bendsøe [10, 11], Bendsøe and Sigmund [13], Borrvall [16], the review articles Eschenauer and Olhoff [33] and Rozvany [80] and the citations in all the mentioned literature. Additionally to the above literature, other topics strongly related to topology optimization can be found, e.g., in Svanberg [103, 104] and Bruyneel, Duysinx, and Fleury [21] concerning convex programming, in Stolpe and Svanberg [101] and Bendsøe and Sigmund [12] about material interpolation schemes and in Sigmund and Petersson [91] and Bourdin [19] about restriction methods. Some applications in electromagnetism are presented by Hoppe, Petrova, and Schulz [42], Yoo and Kikuchi [111]. More theoretical issues are given by Stadler [94] or by Sigmund and Petersson [71].

In the design process the second step after topology optimization is shape optimization, where we tune the shape of the boundary or interfaces. The basic literature on shape optimization is given by Begis and Glowinski [8], Murat and Simon [61], Pironneau [72], Haslinger and Neittaanmäki [39], Haslinger and Mäkinen [38], Sokolowski and Zolesio [93], Börner [15], Delfour and Zolesio [28], Kawohl et al. [46], Mohammadi and Pironneau [60]. Besides the basic textbooks, one can find a lot of theoretical analysis in Bucur and Zolesio [22], Peichl and Ring [67, 68], Petersson and Haslinger [70], Petersson [69]. Papers focused on applications in electromagnetism are, for example, Di Barba et al. [5], Brandstätter et al. [20], Marrocco and Pironneau [58], Takahashi [106].

There are several interesting optimization techniques that have appeared just recently. Within the second SFB period Burger and Mühlhuber [23, 24] solved simultaneously for both the design and state variables, i.e., they minimized at the same time the cost functional as well as the quadratic energy functional of the direct problem. Other papers dealing with this simultaneous approach are Hoppe [41] and Maar and Schulz [57]. This approach results in a saddle point problem. Some solution theory concerning saddle point problems can be found in Schulz [88], Zulehner [113, 114] and in Schöberl and Zulehner [85].

Another challenging issue in optimization is adaptivity. A hierarchical approach in shape optimization was used in the SFB by Lukáš [49, 54]. This approach aims at developing an adaptive optimization method in order to control the error of approximation of the cost functional. We refer to the quite recent papers by Ramm, Maute, and Schwarz [73] and Schleupen, Maute, and Ramm [84] that make use of the

FE–adaptivity in both topology and shape optimization. A multilevel approach was also used in inverse problems, see SCHERZER [83] and KALTENBACHER [45].

1.3 Results and Discussion

In this section we discuss in more details the given problems, our experiences and the obtained results. The first part is concerned with topology and shape optimization problems in structural optimization and magnetostatics, including also our work on geometry and mesh handling strategies. In the last part of this section we discuss several techniques for constructing efficient solver for KKT systems, which arise in all-at-once optimization approaches.

1.3.1 Adaptive Multilevel 3D Topology Optimization Techniques

An Adaptive Multilevel Approach to the Minimal Compliance Problem
Minimizing compliance turned out to be a standard problem in topology optimization. However, it already contains the most basic, but non-trivial difficulties like mesh-dependent solutions, local minima and checkerboard phenomena (see BENDSØE AND SIGMUND [9]. Due to this ill-posedness we need regularization. In our algorithm we combine two filter methods, which are the cornerstones of our adaptive multilevel approach.

An ideal formulation of the minimal compliance problem looks like the following:

$$\ell(\mathbf{u}) = \int_{\Gamma_t} \mathbf{t}\cdot\mathbf{u}\; d\mathbf{s} \;\to\; \min_{\rho,\mathbf{u}} \tag{1a}$$

$$\text{subject to} \qquad a(\rho;\mathbf{u},\mathbf{v}) \;=\; \ell(\mathbf{v}), \qquad \forall\ \mathbf{v}\in V_0, \tag{1b}$$

$$\int_\Omega \rho(\mathbf{x})\; d\mathbf{x} \;\le\; m_0, \tag{1c}$$

$$\rho(\mathbf{x}) \;\in\; \{0,1\}, \qquad \text{a.e. in } \Omega. \tag{1d}$$

In our acutal computations the state equation (1b), describing the corresponding elasticity problem, is formally eliminated and hidden in the objective functional. The energy bilinearform on $V_0\times V_0$ contains a variable material tensor and is given by

$$a(\rho;\mathbf{u},\mathbf{v}) = \int_\Omega \boldsymbol{\varepsilon}\big(\mathbf{u}(\mathbf{x})\big) : \mathbf{C}\big(\rho(\mathbf{x})\big)\boldsymbol{\varepsilon}\big(\mathbf{v}(\mathbf{x})\big)\; d\mathbf{x},$$

where $V_0 = H^1_{\Gamma_u}(\Omega;\mathbf{R}^{\mathbf{d}})$ denotes the set of kinematically admissible displacement fields. The constraint (1d) is relaxed and replaced by $0<\rho_{\min}\le\rho(\mathbf{x})\le 1$, where $\rho_{\min}$ is greater that 0 to assure the elipticity of the bilinearform. In order to penalize intermediate values is to introduce the following nonlinear material tensor:

$$\mathbf{C}_q\big(\rho(\mathbf{x})\big) = \frac{\rho(\mathbf{x})}{1+q\big(1-\rho(\mathbf{x})\big)}\mathbf{C}^0,$$

which is called RAMP (*Rational Approximation of Material Properties*) and was mentioned in Rietz [75] and treated thoroughly in Stolpe and Svanberg [101].

It is well known fact that topology optimization problems lack existance of solutions. Therefore we introduce two filter methods to regularize the problem. The first is called the *mesh-independence filter* and was first proposed in Sigmund [90]. Here, the discrete element sensitivities of an discrete objective $J^h(\boldsymbol{\rho}^h)$ are modified as follows:

$$\frac{\widehat{\partial J^h}}{\partial \rho_k^h} = \frac{1}{\rho_k^h \sum_{i=1}^n H_{i,k}} \sum_{i=1}^n H_{i,k} \rho_i^h \frac{\partial J^h}{\partial \rho_i^h}, \tag{2}$$

where the convolution operator $H_{i,k}$ with filter radius R is defined as

$$H_{i,k} = \max\big\{0, R - \operatorname{dist}(i,k)\big\}, \quad \text{for } i,k = 1,\dots,n.$$

The second filter method is called *Regularized Intermediate Density Control* (*RIDC*) and is discussed in detail in Borrvall and Petersson [17]. Here an additional constraint is added to the optimization problem:

$$P_S(\rho) = \int_\Omega \Big(1 - S\big(\rho(\mathbf{x})\big)\Big)\Big(S\big(\rho(\mathbf{x})\big) - \underline{\rho}\Big)\, d\mathbf{x} \;\le\; \varepsilon_P, \tag{3}$$

where $S : L_2(\Omega) \to L_2(\Omega)$ is an integral operator defined as

$$S(\rho) = \int_\Omega \phi(\mathbf{x},\mathbf{y})\rho(\mathbf{y})\, d\mathbf{y}, \quad \forall\, \mathbf{x} \in \Omega, \tag{4}$$

with the kernel

$$\phi(\mathbf{x},\mathbf{y}) = C(\mathbf{x}) \max\left(0, 1 - \frac{|\mathbf{x}-\mathbf{y}|}{R}\right). \tag{5}$$

Our basic motivation for a multilevel algorithm is to solve the problem efficiently and to save computational costs. This is achieved by solving the problem firstly on a coarse grid to get a first coarse design for rather cheap computational costs. Then we will use this first coarse design as an initial design on a finer grid and repeat the optimization on the finer grid, and so on. Elements inside a region, solely occupied by material or void, far away from the structure's boundary, are very unlikely to be affected by the optimization on finer levels. Far more interesting is the interface between material and void, i.e. the boundary of the structure. It is much more efficient to identify this interface and only refine elements along this interface, instead of an uniform refinement. For identifying the interface the filter operator S, defined in (4), turns out to be a useful tool. We mark the element τ_i to be refined, if

$$\big|\big(\boldsymbol{\Phi}\boldsymbol{\rho}^h\big)_i - \rho_i^h\big| \ge \delta_1 > 0, \tag{6}$$

for some δ_1 with $1 \gg \delta_1 > 0$. In (6) $\boldsymbol{\Phi} \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$ denotes the convolution matrix corresponding to the integral kernel ϕ. In Figure 1 we see an example of the application of

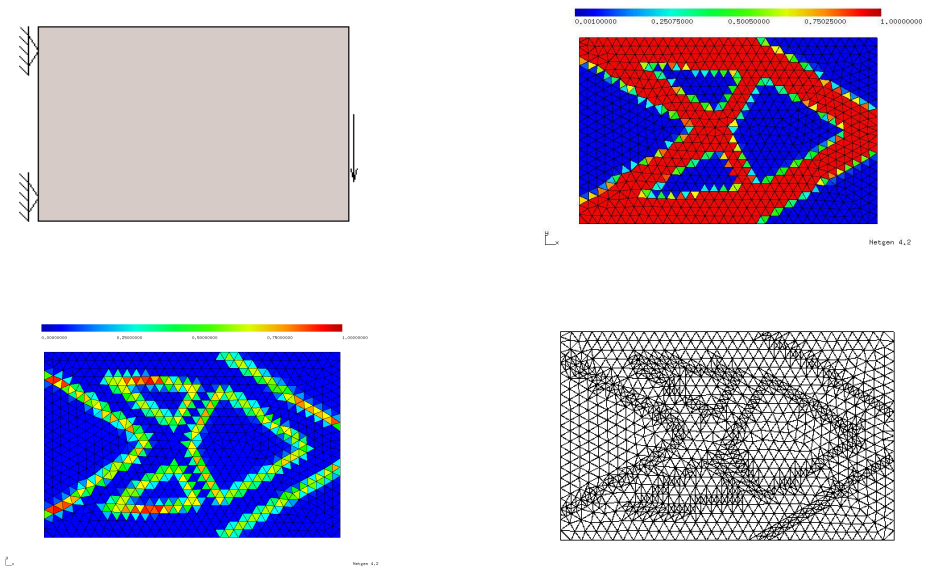

Figure 1: Sketch, coarse solution, identified boundary and refined mesh of the cantilever problem in 2D.

Algorithm 1 An adaptive multilevel approach

Initialize start value $\boldsymbol{\rho}_0^H$, e.g. like $\boldsymbol{\rho}_0^H = m_0/|\Omega|$
Choose the parameters δ_1 and δ_2 with $0 < \delta_1 \ll 1$ and $1 < \delta_2$ respectively.
$l = 0$;

Coarse grid solution $\boldsymbol{\rho}^H$ with MIF and RAMP;
Determine $\boldsymbol{\varepsilon_P^0}$ by $\varepsilon_P^0 = P_S^H(\boldsymbol{\rho}^H)$;
while design not satisfactory **do**
  Mesh-refinement along the interface of void and material. Possible reduction of ε_P:
    $\varepsilon_P^{l+1} = \delta_3 \varepsilon_P^l$, $0 < \delta_3 \leq 1$;
  Fine grid solution $\boldsymbol{\rho}^h$ using RIDC;
  $l = l + 1$;
end while

this refinement idea to the cantilever example in 2D.

The complete algorithm is summarized in Alg. 1. The approach described above was tested with several benchmark examples and we got very good results from all of them. Following a standard finite element procedure the ground structure Ω is partitioned into $n = \mathcal{O}(h^{-d})$ ($n = n_{el} = n_\rho$) triangles τ_i (or tetrahedrons for $d = 3$), where h is the discretization parameter. The density ρ is approximated by a piecewise constant finite element function $\tilde{\rho}$, i.e. $\tilde{\rho}$ is constant over every triangle τ_i. The displacement field $\mathbf{u}$ is

approximated using continuous element-wise quadratic functions. For solving the discrete optimization problems the method of moving asymptotes was used. In Table 1, we list the computational data of the 3D cantilever beam example. For more details we refer

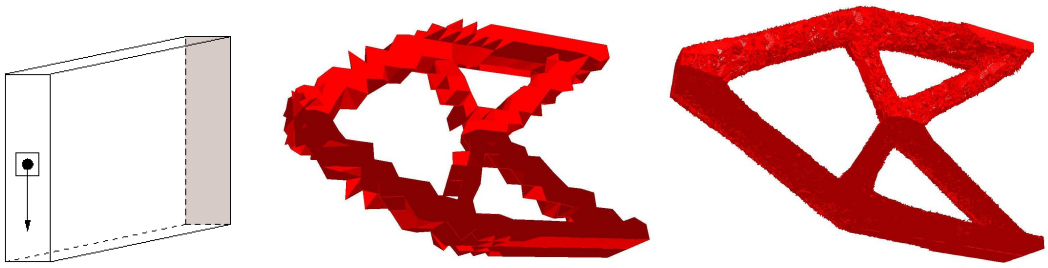

Figure 2: The cantilever beam in 3D: Sketch, coarse grid solution and fine grid solution.

l	N_{el}	$N_{\mathbf{u}}$	t_{state}	t_{∇}	t_{opt}	t_{fil}	t_{it}	*Iter.*
level 0:	1725	9774	1.4	0.2	0.0	0.0	1.7	96
level 1:	8857	41307	11.4	1.2	0.2	0.2	13.4	93
level 2:	49437	214374	76.6	6.7	1.0	10.2	108.5	55
level 3:	189288	794628	330.5	26.1	3.9	160.7	691.8	45

Table 1: Computational features from the 3D cantilever beam example.

to Stainko [96] and to the corresponding chapter in Stainko [98].

Phase-Field Relaxation to Topology Optimization with Local Stress Constraints

In structural optimization there are two design - constraint combinations of particular importance, namely the maximization of material stiffness (minimizing the compliance) at given mass and the minimization of mass while keeping a certain stiffness. The first combination, also known as the minimal compliance problem, seems to be mathematically well understood and various successful numerical techniques to solve the problem have been proposed. The treatment of the second problem is by far less understood and until now there seems to be no approach that is capable of computing reliable (global) optimal designs within reasonable computational effort. The main source of difficulties in this problem is a lack of constraint qualifications for the set of feasible designs, defined by the local stress constraints.

The starting point of our analysis is a reformulation of the equality constraints describing the elastic equilibrium and the local inequality constraints for the stresses into a system of linear inequality constraints as recently proposed by Stolpe and Svanberg [102]. A remaining difficulty is that the arising problem also involves 0-1 constraints

in addition to the linear inequalities. Instead of solving mixed linear programming problems we propose to use a phase–field relaxation of the reformulated problem. Due to the well-known ill-posedness of topology optimization problems we might add a perimeter penalization to the objective functional. The phase–field relaxation consists in using a linear material interpolation function, and additionally, a Cahn-Hillard type penalization functional is used to approximate the perimeter.

Let $\Omega_{\text{mat}} = \{\mathbf{x} \in \Omega \mid \rho(\mathbf{x}) = 1\} \subset \Omega \subset \mathbf{R^d}$ $(d = 2, 3)$, denote the optimal design, which is of course initially unknown. Furthermore, let $\Gamma_{t_0} \subset \Gamma_t$ describe the part of the boundary Γ_t where the traction forces are zero, i.e. $\mathbf{t} = \mathbf{0}$. Then, the stress constrained topology optimization problem that we are going to investigate in this work states as follows:

$$
\begin{aligned}
J(\rho) = \int_\Omega \rho(\mathbf{x})\, d\mathbf{x} \;&\to\; \min_{\rho,\mathbf{u}} && && \text{(7a)}\\
\text{subject to} \qquad \operatorname{div}\boldsymbol{\sigma} \;&=\; 0, && \text{in } \Omega_{\text{mat}}, && \text{(7b)}\\
\boldsymbol{\sigma} - \mathbf{C}\boldsymbol{\varepsilon}(u) \;&=\; \mathbf{0}, && \text{in } \Omega, && \text{(7c)}\\
\mathbf{u} \;&=\; \mathbf{0}, && \text{on } \Gamma_u, && \text{(7d)}\\
\boldsymbol{\sigma}\cdot\mathbf{n} \;&=\; \mathbf{t}, && \text{on } \Gamma_t, && \text{(7e)}\\
\boldsymbol{\sigma}\cdot\mathbf{n} \;&=\; \mathbf{0}, && \text{on } \big(\partial\Omega_{\text{mat}} \setminus \Gamma_t\big) \cup \Gamma_{t_0}, && \text{(7f)}\\
\rho(\mathbf{x}) \;&\in\; \{0,1\}, && \text{a.e. in } \Omega, && \text{(7g)}\\
\Phi^{\min} \;\le\; \Phi\big(\boldsymbol{\sigma}(\mathbf{x})\big) \;&\le\; \Phi^{\max}, && \text{a.e. in } \Omega_{\text{mat}}, && \text{(7h)}\\
\mathbf{u}^{\min} \;\le\; \mathbf{u}(\mathbf{x}) \;&\le\; \mathbf{u}^{\max}, && \text{a.e. in } \Omega. && \text{(7i)}
\end{aligned}
$$

In the bound constraints (7h), Φ denotes a proper stress criterion. For $\Phi(\boldsymbol{\sigma}) = \boldsymbol{\sigma}$ we have that $\boldsymbol{\sigma}^{\min} \le \boldsymbol{\sigma} \le \boldsymbol{\sigma}^{\max}$ and we shall call this criterion total stress. Alternatively, e.g., we can use the von Mises stress criterion, but for sake of simplicity we will use total stresses throughout this report.

For the reformulation of the set of constaints we introduce a $\beta > 0$, such that

$$
\beta\big|\sigma_{ij}(\mathbf{x})\big| \le 1, \qquad \text{a.e. in } \Omega,\ i,j = 1,\dots,d, \tag{8}
$$

and an additional variable $\mathbf{s}$, such that $\mathbf{s}(\mathbf{x}) = \boldsymbol{\sigma}(\mathbf{x})$ if $\rho(\mathbf{x}) = 1$ and $\mathbf{s}(\mathbf{x}) = \mathbf{0}$ if $\rho(\mathbf{x}) = 0$,

i.e. $\mathbf{s} = \rho\boldsymbol{\sigma}$. Then the equivalent reformulation of the set of constraints looks like:

$$\begin{aligned}
\operatorname{div}\mathbf{s} &= 0, && \text{in } \Omega, && (9a)\\
\boldsymbol{\sigma} - \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u}) &= \mathbf{0}, && \text{in } \Omega, && (9b)\\
\mathbf{u} &= \mathbf{0}, && \text{on } \Gamma_u, && (9c)\\
\mathbf{s}\cdot\mathbf{n} &= \mathbf{t}, && \text{on } \Gamma_t, && (9d)\\
\mathbf{s}\cdot\mathbf{n} &= \mathbf{0}, && \text{on } \Gamma_{t_0}, && (9e)\\
-(1-\rho)\mathbf{1} \le \beta(\boldsymbol{\sigma}-\mathbf{s}) &\le (1-\rho)\mathbf{1}, && \text{in } \Omega, && (9f)\\
\rho(\mathbf{x}) &\in \{0,1\}, && \text{a.e. in } \Omega, && (9g)\\
\rho(\mathbf{x})\boldsymbol{\sigma}^{\min} \le \mathbf{s}(\mathbf{x}) &\le \rho(\mathbf{x})\boldsymbol{\sigma}^{\max}, && \text{a.e. in } \Omega, && (9h)\\
\mathbf{u}^{\min} \le \mathbf{u}(\mathbf{x}) &\le \mathbf{u}^{\max}, && \text{a.e. in } \Omega. && (9i)
\end{aligned}$$

All the constraints in (9) are linear with respect to the vector of unknowns $(\rho, \mathbf{u}, \boldsymbol{\sigma}, \mathbf{s})$, except for $\rho(\mathbf{x}) \in \{0,1\}$ almost everywhere in Ω. We now replace the 0-1 constraint $\rho(\mathbf{x}) \in \{0,1\}$ by the following continuous version $\rho(\mathbf{x}) \in [0,1]$. Moreover, we approximate a perimeter term by the Cahn-Hilliard term and add it to the objective:

$$J_\epsilon(\rho) = \gamma \int_\Omega \rho(\mathbf{x})\ d\mathbf{x} + \frac{\epsilon}{2}\int_\Omega \left|\nabla\rho(\mathbf{x})\right|^2\ d\mathbf{x} + \frac{1}{\epsilon}\int_\Omega W\big(\rho(\mathbf{x})\big)\ d\mathbf{x}. \tag{10}$$

The term $\int_\Omega W(\rho(\mathbf{x})\ d\mathbf{x}$ favorites those designs which take values close to 0 or 1 (*phase separation*), while the term $\int_\Omega |\nabla\rho(\mathbf{x})|^2\ d\mathbf{x}$ penalizes the spatial inhomogeneity of ρ. The theorem of Modica and Mortola tells that the minimizers of (10) converge to the minimizers of $\int_\Omega \rho(\mathbf{x})\ d\mathbf{x}$ in the sense of Γ-convergence (see Modica and Mortola [59]). The resulting relaxed paremeter dependendent problem is now given by the objective functional (10) and by the constraints (9), where (9g) is replaced by $0 \le \rho(\mathbf{x}) \le 1$. The problem is now solved for a decreasing sequence of the parameter $\varepsilon \to 0$. For the relaxed problem it is now possible to show the existance of solutions in the corresponding set of feasible designs.

After a standard finite element discretization we end up with a large scale optimization problem, that now fulfills constraint qualifications. We solved the discrete optimization problems using *Ipopt*, which is a free available optimization code realizing a primal-dual interior-point optimization method (see Wächter et al [109]). This new approach to stress constrained topology optimization resulted in a joint paper of Burger and Stainko [26].

An Optimal Solver to a KKT-System

Over the last two decades interior-point methods turned out to be efficient optimization methods for solving large-scale nonlinear optimization problems. Most of the computing time is actually spent to the solution of linear systems arising from the linearization of the primal-dual optimality conditions. Instead of solving the nonsymmetric systems, a

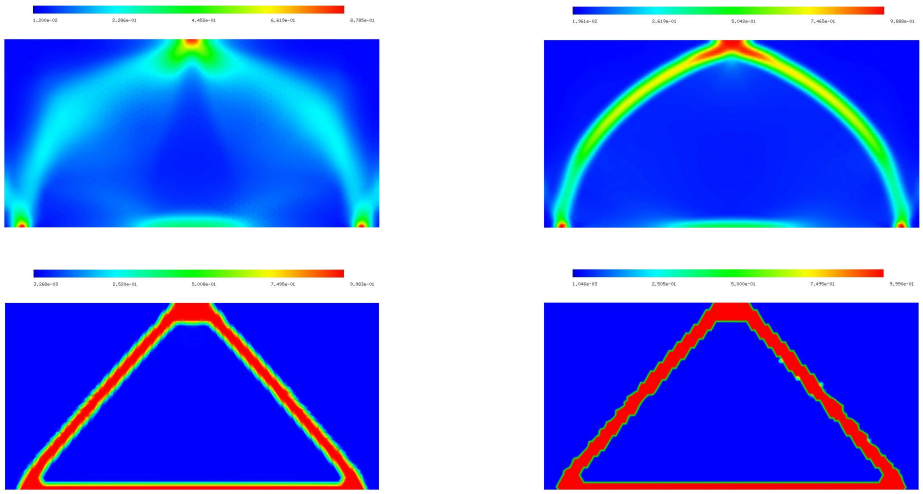

Figure 3: Optimal designs of the 4 level ϵ-continuation with $\epsilon_0 = 0.1$, $\epsilon_1 = 0.05$, $\epsilon_2 = 0.025$, and $\epsilon_3 = 0.0125$, respectively.

symmetric system like

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \triangle \mathbf{x} \\ \triangle \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \tag{11}$$

can be achieved by some elimination steps. Multigrid methods certainly belong to the most efficient methods for solving large-scale systems, arising from discretized partial differential equations. While the construction of such methods for symmetric and positive definite systems is quite standard, this is not the case for saddle point problems. A successful construction of a solver with optimal complexity for linear systems like (11) would yield a significant speedup for an interior-point method. One of the most important ingredients of an efficient multigrid method is an appropriate smoother, i.e. a simple iterative smoothing procedure. Here we consider a multiplicative Schwarz-type iteration method as a smoother in a multigrid method. Each iteration step of such a multiplicative Schwarz-type smoother consists of the solution of several small local saddle point problems, i.e. small local versions of the problem (11).

After several eliminiation steps, the optimality system of the interior-point formulation of the stress constrained topology optimization problem turns into the following symmetric saddle point problem

$$\begin{pmatrix} \boldsymbol{\mathcal{K}}_{\rho\rho} & \boldsymbol{\mathcal{K}}_{\rho u} & \boldsymbol{\mathcal{K}}_{\rho s} & \mathbf{0} \\ \boldsymbol{\mathcal{K}}_{\rho u}^T & \boldsymbol{\mathcal{K}}_{uu} & \boldsymbol{\mathcal{K}}_{us} & \mathbf{0} \\ \boldsymbol{\mathcal{K}}_{\rho s}^T & \boldsymbol{\mathcal{K}}_{us}^T & \boldsymbol{\mathcal{K}}_{ss} & \mathbf{D}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{D} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \triangle \boldsymbol{\rho}^h \\ \triangle \mathbf{u}^h \\ \triangle \mathbf{s}^h \\ \triangle \boldsymbol{\lambda}_0^h \end{pmatrix} = \mathbf{f}^h, \tag{12}$$

where the block matrices of the coefficient matrix are combinations of the finite element matrices from the finite element formulation of the primal-dual optimality conditions. We

start the introduction of the multiplicative Schwartz- type smoother with a decomposition of the spaces

$$V = \sum_{i=1}^{l} V^i \qquad \text{and} \qquad Q = \sum_{i=1}^{l} Q^i.$$

Moreover, we have to introduce linear operators for each subspace to set up the local sub-problems:

$$\mathbf{P}_{Vi} : \mathbf{R^{n_i}} \to \mathbf{R^n} \qquad \text{and} \qquad \mathbf{P_{Q_i}} : \mathbf{R^{m_i}} \to \mathbf{R^m}, \qquad \text{for } \mathbf{i = 1, \ldots, l}, \tag{13}$$

with n_i, m_i denoting the dimensions of the local subspaces V_i and Q_i, respectively. The matrices $\mathbf{P}_{Vi}$ and $\mathbf{P}_{Q_i}$ denote prolongation operators with the associated restriction operators $\mathbf{P}_{V_i}^T$ and $\mathbf{P}_{Q_i}^T$, respectively. Now we define the *multiplicative Schwarz smoother* based on the above subspace decomposition as the following procedure: Set $\mathbf{w}^0 = \mathbf{0}$ and $\mathbf{r}^0 = \mathbf{0}$ and compute

$$\begin{pmatrix} \mathbf{w}^i \\ \mathbf{r}^i \end{pmatrix} = \begin{pmatrix} \mathbf{w}^{i-1} \\ \mathbf{r}^{i-1} \end{pmatrix} + \mathbf{P}_i \hat{\boldsymbol{\mathcal{K}}}_i^{-1} \mathbf{P}_i^T \left(\begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} - \boldsymbol{\mathcal{K}} \begin{pmatrix} \mathbf{w}^{i-1} \\ \mathbf{r}^{i-1} \end{pmatrix} \right), \qquad \text{for } i = 1, \ldots, l, \tag{14}$$

where $\hat{\boldsymbol{\mathcal{K}}}_i$ denotes the coefficient matrix of the local saddle point problem. Finally we define the multiplicative smoother as

$$\boldsymbol{\mathcal{S}}_M \begin{pmatrix} \mathbf{u}^k \\ \mathbf{p}^k \end{pmatrix} = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{p}^k \end{pmatrix} + \begin{pmatrix} \mathbf{w}^l \\ \mathbf{r}^l \end{pmatrix}. \tag{15}$$

We refer to Schöberl and Zulehner [85] for a theoretical analysis for the convergence and smoothing properties of the additive smoother.

For the numerical results we choose $\Omega = (0,1) \times (0,1)$ and decompose it into a regular triangulation $\mathcal{T}_h^k = \{\tau_i \mid i = 1, \ldots, n_k\}$ for each level k of a hierarchy of l nested meshes with $3 \leq k \leq l$. That means that level $k = 3$ is the coarsest grid where the corresponding linear system is solved exactly. For each level k we assemble the block matrices that finally build up the saddle point system (12). In order to test the multiplicative patch smoother (14) - (15) we solved the saddle point system (12) on a hierarchy with an increasing number of meshes. We set $\mathbf{f}_k = \mathbf{0}$ and used randomly chosen starting values for $\triangle \mathbf{x}_k^0$ for the exact solutions $\triangle \mathbf{x}_k$. For constructing the local subproblems we decomposed the grid $\mathcal{T}_h^k$ into m_k overlapping patches, where m_k denotes the number of nodes on level k. Each patch consists of the at most 6 surrounding triangles for each node. We approximated the density ρ, the displacements $\mathbf{u}$, and the Lagrangian multiplier $\boldsymbol{\lambda}_0$ with linear elements and the stresses $\mathbf{s}$ with constant elements. The corresponding subspaces V_i, for $i = 1, \ldots, m_k$, consist now of the degrees of freedom of the node i, related to the approximations of the density and the displacement components, and the degrees of freedom in the surrounding elements, related to the stress components. The subspaces Q_i, for $i = 1, \ldots, m_k$, consist of the unknowns at node i with respect to the approximation of the Lagrangian multiplier $\boldsymbol{\lambda}_0$. Figure 4 shows an example of a patch, where the places marked with a '■' indicate

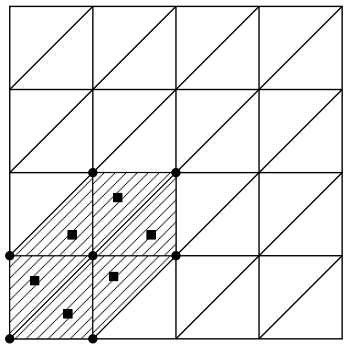

Figure 4: Patch of a local saddle point problem.

the unknowns of the constant elements and the places marked with a '$\bullet$' indicate the unknowns of the linear elements. For the actual numerical tests we used a W-cycle with s smoothing steps ($s/2$ pre- and $s/2$ post-smoothing step. We stopped the iteration process when the initial defect was reduced by a factor of 10^{-8}, measured by the Euclidean norm. In Table 2 we list the convergence data for the following choice of parameters: $\epsilon = 10^{-4}$, $\mu = 10^{-6}$, and $\boldsymbol{\nu}_i^h = \mathbf{1}$ for $i = 1, \ldots, 8$. The table shows the typical multigrid convergence behavior, i.e., convergence rates that are asymptotic independent of the grid level and an asymptotic constant number of iterations. A more detailed description of this solution

Level	Unknowns	Smoothing steps			
		2		4	
		Iterations	Conv. Factor	Iterations	Conv. Factor
4	725	39	0.621	19	0.376
5	2853	25	0.478	14	0.258
6	11333	24	0.460	13	0.226
7	45189	22	0.427	12	0.210
8	180485	22	0.425	12	0.211

Table 2: Convergence rates for a W-cycle and an error reduction by a factor of 10^{-8} ($\epsilon = 10^{-4}$, $\mu = 10^{-6}$, $\boldsymbol{\nu}_i^h = \mathbf{1}$ for $i = 1, \ldots, 8$).

approach can be found in STAINKO [99].

Topology optimization for magnetostatics: setting
We treated topology optimization governed by a nonlinear magnetostatic problem. Let us consider a fixed computational domain $\Omega \subset \mathbb{R}^d$, where $d = 2, 3$. Let $\Omega_\mathrm{d} \subset \Omega$ be the subdomain where the designed structure can arise. The set of admissible material distributions is denoted by $\mathcal{Q} := \{\rho \in L^2(\Omega_\mathrm{d}) \mid 0 \leq \rho \leq 1\}$. We need to penalize the

intermediate values. To this end we introduce the following penalization:

$$\widetilde{\rho}_p(\rho) := \frac{1}{2}\left(1 + \frac{1}{\arctan(p)}\arctan(p(2\rho-1))\right), \quad p > 0$$

which, unlike SIMP nor RAMP, penalizes 0 and 1 equally, see also Fig. 5. Further, we

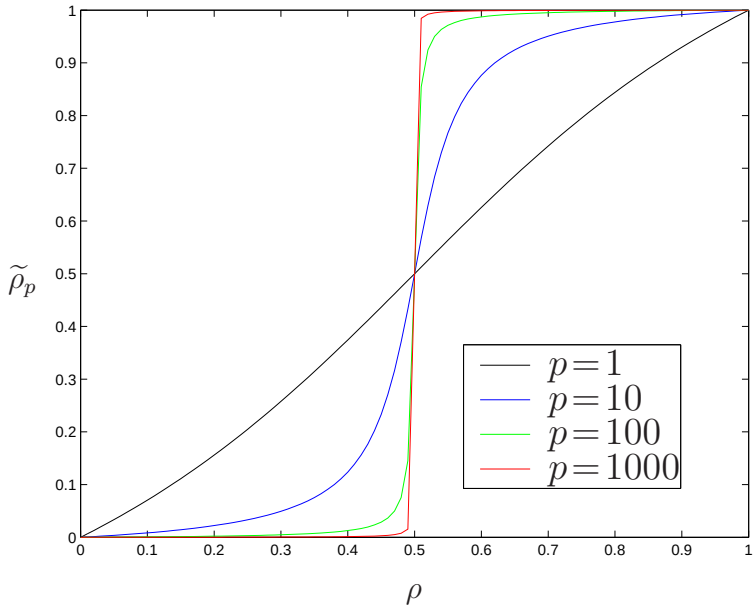

Figure 5: Penalization of intermediate values

consider the following nonlinear magnetic reluctivity:

$$\nu(\eta, \widetilde{\rho}) := \begin{cases} \nu_0 + (\nu(\eta) - \nu_0)\widetilde{\rho}, & \text{in } \Omega_{\mathrm{d}} \\ \nu_0, & \text{otherwise,} \end{cases}$$

where $\nu(\eta) := \nu_1 + (\nu_0 - \nu_1)\frac{\eta^8}{\eta^8 + \nu_0^{-1}}$ is due to Křížek and Neittaanmäki [48], p. 134, and ν_0, ν_1 are the reluctivities of the air and ferromagnetics, respectively. Finally, we consider a cost functional $\mathcal{I} : \mathbf{L}^2(\Omega) \times \mathcal{Q} \mapsto \mathbb{R}$, possibly involving penalization of state constraints. Given a maximal volume $V_{\max}$ of the designed structure, the 3D topology optimization problem governed by the nonlinear magnetostatics then reads as follows:

$$\begin{cases} & \min_{\rho \in \mathcal{Q}} \mathcal{I}(\mathbf{curl}(\mathbf{u}), \widetilde{\rho}(\rho)) \\ \text{w.r.t.} & \\ & \int\limits_{\Omega_{\mathrm{d}}} \widetilde{\rho}(\rho)\, d\mathbf{x} \leq V_{\max} \\ & \int\limits_{\Omega} \nu\left(\|\mathbf{curl}(\mathbf{u})\|, \widetilde{\rho}(\rho)\right) \mathbf{curl}(\mathbf{u}) \cdot \mathbf{curl}(\mathbf{v})\, d\mathbf{x} = \int\limits_{\Omega} \mathbf{J} \cdot \mathbf{v}\, d\mathbf{x} \text{ in } \mathbf{H}_{0,\perp}(\mathbf{curl}; \Omega), \end{cases}$$

where $\mathbf{J} \in \mathbf{L}^2(\Omega)$ is a divergence–free current density and where the ansatz space $\mathbf{H}_{0,\perp}(\mathbf{curl};\Omega)$ contains such functions $\mathbf{v} \in \mathbf{L}^2(\Omega)$ that, in a weak sense, $\mathbf{curl}(\mathbf{v}) \in \mathbf{L}^2(\Omega)$ and $\mathbf{n} \times \mathbf{v} = \mathbf{0}$ along $\partial\Omega$ and that are additionaly $\mathbf{L}^2$–orthogonal to the kernel of $\mathbf{curl}$.

For the numerical solution, the 3D problem is discretized by the finite element method using the lowest order edge Nédélec elements on tetrahedra, while we use the lowest order nodal Langrange elements on triangles in case of the 2D reduced problem. The design material distribution is elementwise constant. In Figure 6 there are 2D (a quarter of the geometry) and 3D (an eigth of the geometry) optimal designs depicted for the electromagnet benchmark problem described below. The 2D problem was solved for 55104 design and 66877 state variables using a multilevel method described below, as well. The 3D problem was solved for 14000 design variables and 33323 state ones. Let us note that we have not employed any regularization technique, however, referring to Figure 6 (left) it seems to be useful.

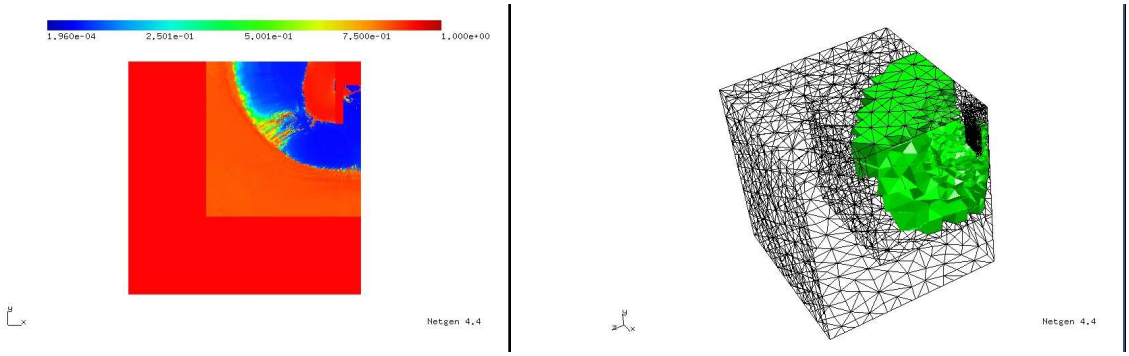

Figure 6: Optimal 2D and 3D shapes

Topology optimization for magnetostatics: sensitivity analysis

The outer optimization problem is solved using the steepest descent method whereas the nonlinear state problem is eliminated by means of nested Newton iterations with a bisection line–search. We need to provide the gradient of the state solution $\mathbf{u}$ w.r.t. the elementwise constant design material function ρ. To this end we differentiate (by hand) the state solution procedure described in Algorithm 2There we denote by $\boldsymbol{A}^{\text{nonlinear}}(\boldsymbol{u},\boldsymbol{\rho})$ and $\boldsymbol{A}^{\text{linear}}(\boldsymbol{\rho})$ the stiffness matrices assembled for the design $\boldsymbol{\rho}$, while in the latter case the linear reluctivity is considered, i.e., $\nu^{\text{linear}}(\eta) := \eta$. Let further $\boldsymbol{f}$ denote the assembled right hand side of the state problem and $\boldsymbol{u}$, $\boldsymbol{\rho}$, and I the vector counterparts of $\mathbf{u}$, ρ, and $\mathcal{I}$, respectively.

In Figure 7 there is the nonlinear reluctivity depicted. We can see that only the pole heads, as being close to the coils, behave nonlinearly. Since we are computing the cost functional in the air, the nonlinearities don't have any influence to the resulting optimal design.

Let us now consider the linear magnetostatic state problem. We use a coupling of the

Algorithm 2 Solving the nonlinear state problem

Given $\boldsymbol{\rho}$
Solve $\boldsymbol{A}^{\text{linear}}(\boldsymbol{\rho}) \cdot \boldsymbol{u}^0 = \boldsymbol{f}$
Assemble $\boldsymbol{f}^0 := \boldsymbol{f} - \boldsymbol{A}^{\text{nonlinear}}(\boldsymbol{u}^0, \boldsymbol{\rho})$
for $i = 1, \ldots, k$ **do**
  Solve $\boldsymbol{A}^{\text{nonlinear}'}_{u}(\boldsymbol{u}^{i-1}, \boldsymbol{\rho}) \cdot \boldsymbol{w}^i = \boldsymbol{f}^{i-1}$
  Line search $\tau^i := \text{argmin}_\tau \left\| \boldsymbol{f} - \boldsymbol{A}^{\text{nonlinear}}(\boldsymbol{u}^{i-1} + \tau \boldsymbol{w}^i, \boldsymbol{\rho}) \right\|$
  $\boldsymbol{u}^i := \boldsymbol{u}^{i-1} + \tau^i \boldsymbol{w}^i$
  $\boldsymbol{f}^i := \boldsymbol{f} - \boldsymbol{A}^{\text{nonlinear}}(\boldsymbol{u}^i, \boldsymbol{\rho})$
  Store $\boldsymbol{w}^i$ and τ^i
end for
Store $\boldsymbol{u}^k$
Calculate objective $I(\boldsymbol{u}^k, \boldsymbol{\rho})$

Algorithm 3 Adjoint Newton method for the nonlinear state problem

Given $\boldsymbol{\rho}$, $\boldsymbol{u}^k$, $\{\boldsymbol{w}^i\}_{i=1}^k$ and $\{\tau^i\}_{i=1}^k$ stored in the previous call of the Newton method
$\boldsymbol{\lambda} := I'_{\boldsymbol{u}}(\boldsymbol{u}^k, \boldsymbol{\rho})$
$\boldsymbol{\omega}^{\text{nonlinear}} := \boldsymbol{0}$
for $i = k, \ldots, 1$ **do**
  $\boldsymbol{u}^{i-1} := \boldsymbol{u}^i - \tau^i \boldsymbol{w}^i$
  Solve $\boldsymbol{A}^{\text{nonlinear}'}_{u}(\boldsymbol{u}^{i-1}, \boldsymbol{\rho})^T \cdot \boldsymbol{\eta} = \boldsymbol{\lambda}$
  Assemble $\boldsymbol{\omega}^{\text{nonlinear}} := \boldsymbol{\omega}^{\text{nonlinear}} + \tau^i \boldsymbol{A}^{\text{nonlinear}''}_{u\rho}(\boldsymbol{u}^{i-1}, \boldsymbol{w}^i, \boldsymbol{\rho})^T \cdot \boldsymbol{\eta}$
  Assemble $\boldsymbol{\lambda} := \boldsymbol{\lambda} + \tau^i \boldsymbol{A}^{\text{nonlinear}''}_{uu}(\boldsymbol{u}^{i-1}, \boldsymbol{w}^i, \boldsymbol{\rho})^T \cdot \boldsymbol{\eta}$
end for
Solve $\boldsymbol{A}^{\text{linear}}(\boldsymbol{\rho})^T \cdot \boldsymbol{\eta} = \boldsymbol{\lambda}$
Assemble $\boldsymbol{\omega}^{\text{linear}} := \boldsymbol{A}^{\text{linear}'}_{\rho}(\boldsymbol{u}^0, \boldsymbol{\rho})^T \cdot \boldsymbol{\eta}$
Calculate the gradient of the objective $I'_{\boldsymbol{\rho}}(\boldsymbol{u}^k(\boldsymbol{\rho}), \boldsymbol{\rho}) := \boldsymbol{\omega}^{\text{nonlinear}} + \boldsymbol{\omega}^{\text{linear}}$

outer steepest–descent optimization iterations with the nested multigrid preconditioned conjugate gradient (PCG) method. The idea is to use the information about the coarsely optimized design as well as the coarse grid preconditioner, see Algorithm 3

Unlike the minimal compliance problem, here the design hardly changes when starting too far from the intermediate value 0.5, see the locking effect in Figure 8. Therefore, one has to shrink the optimal coarse design to a small interval close to 0.5 and use this as the initial guess at the next level. Unfortunately, it makes the coarse preconditioner useless in some first iterations of the steepest descent at this actual level. The algorithm is then still quite effective in 2D, see Table 3, however, we have not managed to do so in 3D. This motivates us to focus our effort on applying shape optimization after a coarsely optimized topology design is available.

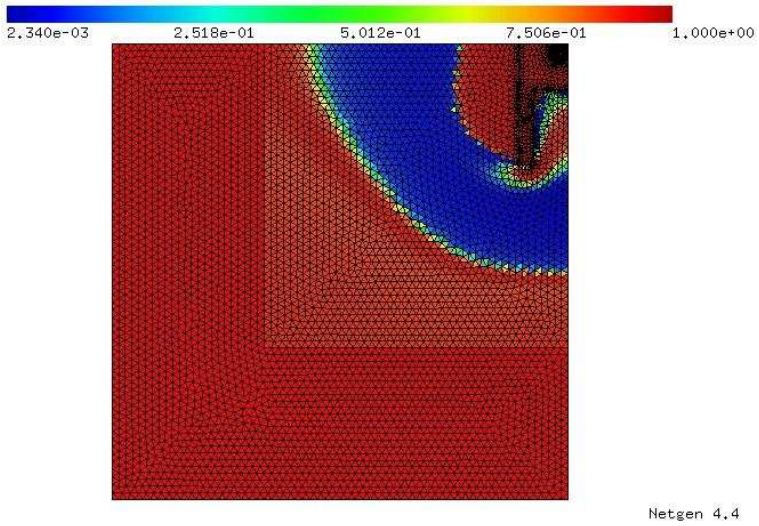

Figure 7: Nonlinear magnetic reluctivity

Algorithm 4 Steepest descent iterations coupled with nested multigrid PCG

Discretize at the first level $\rightsquigarrow h^1, \boldsymbol{\rho}^1_{\text{init}}, \boldsymbol{A}^1(\boldsymbol{\rho}^1_{\text{init}})$
Solve by the steepest–descent method and the nested direct solver $\rightsquigarrow \boldsymbol{\rho}^1_{\text{opt}}$
Store the first level preconditioner $\boldsymbol{C}^1_{\text{opt}} := \boldsymbol{A}^1(\boldsymbol{\rho}^1_{\text{opt}})^{-1}$
for $l = 2, \ldots$ **do**
  Refine $h^{l-1} \rightsquigarrow h^l$
  Increase the penalty
  Prolong $\boldsymbol{\rho}^{l-1}_{\text{opt}} \rightsquigarrow \boldsymbol{\rho}^l_{\text{init}}$
  Solve by the steepest–descent method and the nested multigrid solver $\rightsquigarrow \boldsymbol{\rho}^l_{\text{opt}}$
  Store the l–th level preconditioner $\boldsymbol{C}^l_{\text{opt}}$
end for

1.3.2 Geometry Handling

The main issue of this topic is to find a proper hirarchical geometric representation and related numerical techniques for dealing with shapes of the structures which arise from topology optimization. Here we mainly cooperate with the subproject F1315. So far we had several meetings with Bert Jüttler, Mohamed Shallaby, and Pavel Chalmoviansky and discussed preliminary results. A joint paper with Pavel Chalmoviansky is in preparation.

Implicit shape representation: B–spline wavelets
First, we attempted to use implicit representation of the shapes which gives us a strong connection between topology and shape optimization as well as to the level–set methods. The research was initiated by Ph.D. thesis Shallaby [89]. He provides techniques for implicitization of shapes using tensor–product B–splines. Moreover, a hierarchical

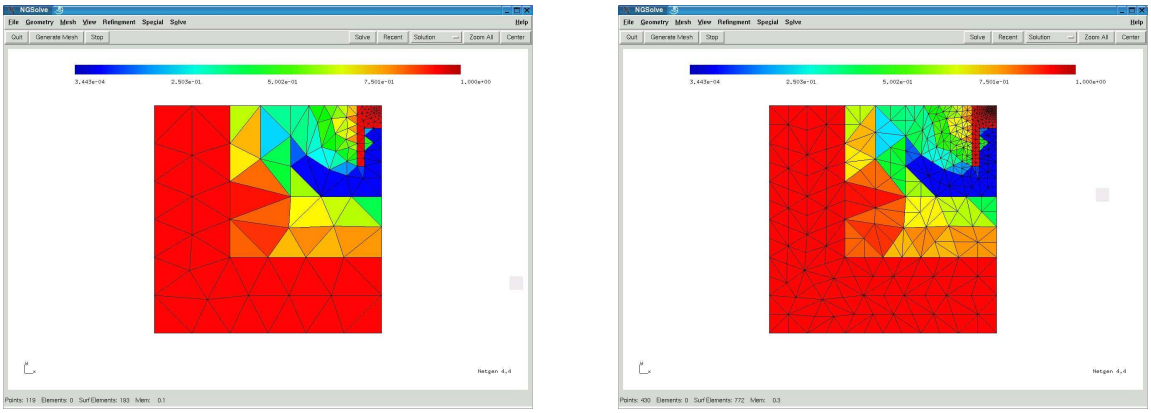

Figure 8: Locking in the coarsely optimized design

level	design vars.	state vars.	steepest desc. iters.	PCG iters. (rel. prec. 10^{-2})	total time
1	861	1105	7		2s
2	3444	4282	6	3	9s
3	13776	16855	13	3	80s
4	55104	66877	40	5–10	29min 17s

Table 3: Steepest descent coupled with multigrid for 2D topology optimization

representation of complex 2D geometries is supported by means of wavelet techniques. In Shallaby [89] he presents a hierarchical construction which starts from the finest geometry and the details (wavelet coefficients) are neglected at coarser levels. Still, since within our multilevel framework we proceed from the coarse to finest geometry, one has to develop the method the other way round. One can use a p–refinement of the B–splines, see Figure 9, however, the hierarchy then is not nested. Therefore, we prefer an h–refinement instead using a lower order, e.g., bilinear B–splines.

It is natural to use this representation with a level set type method. After several meetings with Martin Burger, we started to work on using the phase field method. In particular, we only added the following phase field penalization term to the objective:

$$\mathcal{P}_p(\rho) := p \int_{\Omega_d} \rho^2(1-\rho^2)\, d\mathbf{x}.$$

This term penalizes the intermediate values instead of the arctan–like penalization $\widetilde{\rho}_p(\rho)$. However, this method had taken too long before the convergence was achieved. Typically, after 800 iterations the method was still in progress with the intermediate results depicted in Figure 10. A similar numerical evidence was also observed by Martin Burger. Due to that our work concerning phase–field method as well as implicit shape representations

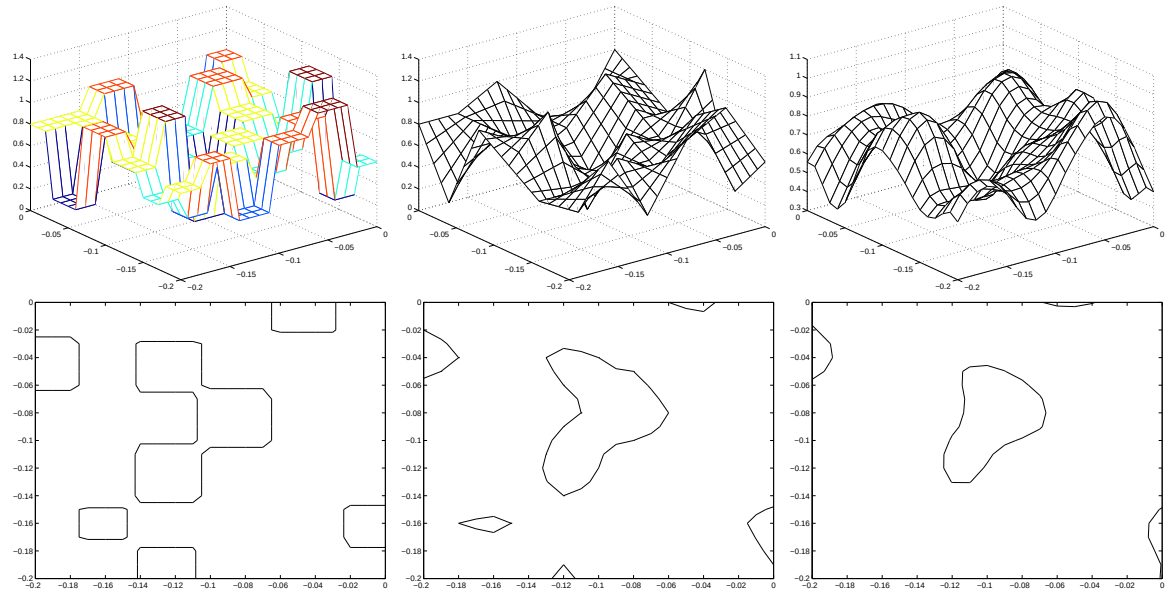

Figure 9: B–spline implicit shape representation of the order 0, 1, 2 and their 0 level sets

has stalled.

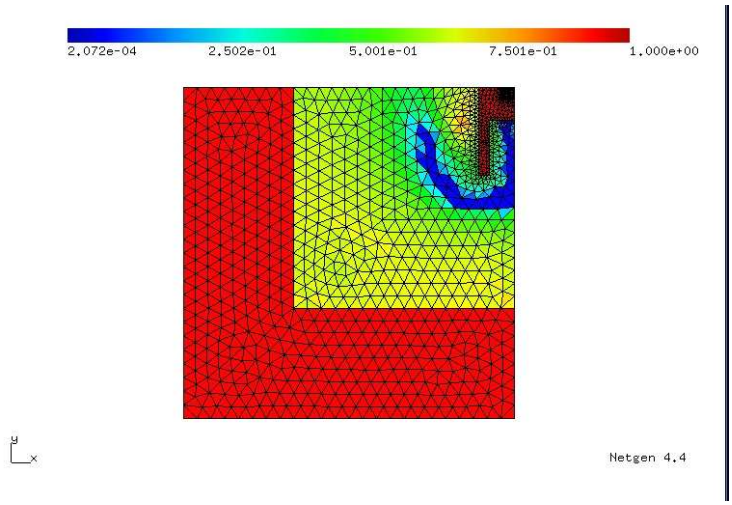

Figure 10: Design development after 800 iterations of the phase field method

Parameteric shape representation using Bezier curves
Next, we discussed the geometric issues with Pavel Chalmoviansky and we realized that Béziér curves or surfaces have two nice properties that make them superior to be used. First, one can introduce new control nodes so that the Béziér curve or surface is the same,

i.e., we have a nested hierarchy of them. Second, just a few such refinements makes the control polygon close enough to the Béziér shape itself. Therefore, it is sufficient to work with polygons only. In Figure 11 there are 3D pole heads of the electromagnet modelled by the Bézier surfaces.

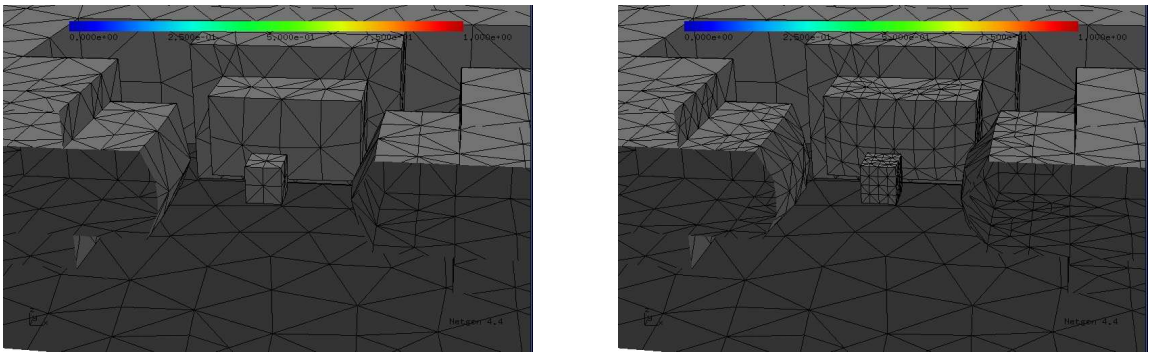

Figure 11: Pole heads modelled by Bézier surfaces in a nested hierarchy

Approximation of shapes arising from 2D topology optimization
Another issue that have been tackled is an integration of topology and shape optimization. As we suggested in the Proposal 2003, first we want to solve a rather coarsely discretized topology optimization problem, then, we identify the boundary and/or interfaces of the resulting structure, approximate the shapes using the Bézier parameterization, and finally, proceed on with the multilevel shape optimization.

In particular, so far we have tested an approximation of a 2D single topology. Given a coarsely optimized material distribution function $\rho^h \in \mathcal{Q}$ and given a number of sets of admissible shapes $\mathcal{U}_1^h, \ldots, \mathcal{U}_n^h$, we are looking for the optimal Bézier parameters by means of the least squares:

$$\min_{\alpha_i^h \in \mathcal{U}_i^h} \int_{\Omega_\mathrm{d}} \left(\rho^h - \chi\left(\Omega_1(\alpha_1^h, \ldots, \alpha_n^h)\right)\right)^2 \, d\mathbf{x},$$

where $\Omega_1(\alpha_1^h, \ldots, \alpha_n^h)$ is the domain occupied with the ferromagnetics (the reluctivity ν_1) the boundary of which is controlled by the shapes, and where $\chi : \mathbb{R}^d \to \{0, 1\}$ is its characteristic function. The result is depicted in Figure 12. For more details we refer to Chalmoviansky and Lukas [55]. Note that a similar issue has recently appear in F1306 project, see Nübel, Düster, and Rank [63]

1.3.3 Adaptive Multilevel 3D Shape Optimization Techniques

Let $\alpha \in \mathcal{U}$ describe the boundary of the ferromagnetic domain Ω_1 so that $\Omega_0 := \Omega \setminus \Omega_1$ is the air domain. We consider the following shape optimization problem governed with

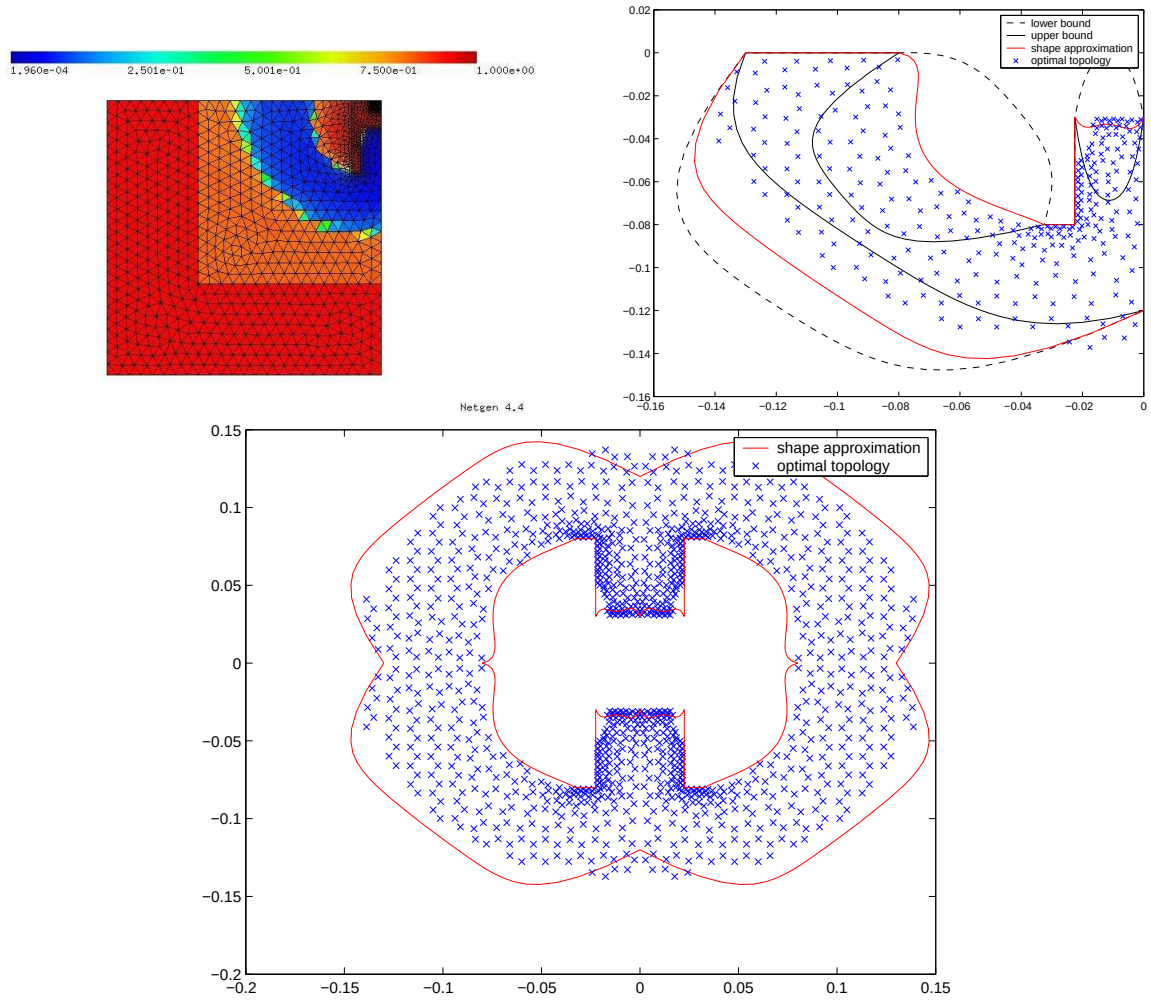

Figure 12: Approximation of shapes arising from 2D topology optimization

the 3D linear magnetostatics:

$$\begin{cases} \min\limits_{\alpha\in\mathcal{U}} \mathcal{I}(\mathbf{curl}(\mathbf{u}),\alpha) \\ \text{w.r.t.} \\ \int\limits_{\Omega_0(\alpha)} \nu_0 \mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{v})\,d\mathbf{x} + \int\limits_{\Omega_1(\alpha)} \nu_1 \mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{v})\,d\mathbf{x} = \\ = \int\limits_{\Omega} \mathbf{J}\cdot\mathbf{v}\,d\mathbf{x} \text{ in } \mathbf{H}_{0,\perp}(\mathbf{curl};\Omega), \end{cases}$$

where all the remaining symbols have the same meaning as in the case of topology optimization in magnetostatics. The problem is discretized by means of the finite element method. The mesh is deformed in accordance to the shape changes by means of solution to an artificial discretized linear elasticity problem. The 2D and 3D results are depicted

level	design vars.	state vars.	Newton iters.	PCG iters. (rel. prec. 10^{-2})	total time
1	4	17653	2	3	4min
2	16	34750	2	3	8min
3	64	93265	2	3	1h 7min

Table 4: Newton method coupled with multigrid for 3D shape optimization

in Figure 13.

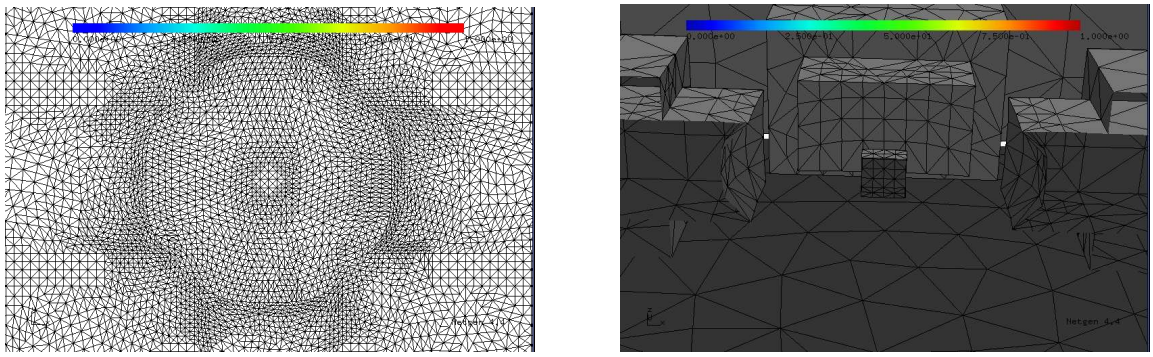

Figure 13: Optimal 2D and 3D shapes of the pole heads of the electromagnet

Newton method coupled with geometric multigrid
The algorithm is similar to Algorithm 4while now we use the Newton outer iterations. The numerical performance for the 3D problem can be seen in Table 4. Note that there was also a 0–th level, where we only assembled the coarsest grid preconditioner for the initial design which is then efficiently used by the multigrid.

1.3.4 Multigrid Methods for KKT Systems

Another part of our investigation was concerned with multigrid methods for solving large-scale systems of discretized mixed variational problems. The main applications considered here are optimization problems in function spaces with constraints in form of partial differential equations (PDEs). The necessary first-order optimality conditions on a solution of such a problem can be written as a mixed variational problem, usually called the optimality system or Karush-Kuhn-Tucker (KKT) system.

In particular, we considered elliptic optimal control problems. In such problems the primal unknown, say x, consists of two parts: a function y, the so-called state, and a function u, the so-called control. The problem is to find $x = (y, u)$ from appropriate function spaces that minimizes a given cost functional subject to a constraint, the so-called state equation, which, for each control u, is an elliptic boundary value problem in y.

The corresponding KKT system involves another (dual) unknown, say p (the Lagrangian multiplier or the adjoint state), and consists of three components: the state equation, the adjoint state equation, which, for each state y, is an elliptic boundary value problem in p, and the control equation, which is typically an algebraic relation between u and p.

In principle, there are two different approaches to take advantage of the multigrid idea. One way is to use an outer iteration, typically a preconditioned Richardson method (possibly accelerated by a Krylov subspace method), applied to the discretized problem. For typical preconditioners of KKT systems in elliptic optimal control, see, e.g., Batterman and Heinkenschloss [6], Battermann and Sachs [7], Biros and Ghattas [14] and Hazra and Schulz [40]. These preconditioners usually rely on efficient solvers or preconditioners for the underlying PDEs and on the construction of a good preconditioner for the corresponding Schur complement, which is the reduced Hessian of the Lagrangian. Multigrid techniques (as an inner iteration) can be used for (some or all of) these components, see, e.g., Dreyer, Maar, and Schulz [30], Hackbusch [36]. We recently contributed to this approach, see Schöberl and Zulehner [87] and the description at the end of this section.

The other way is to use multigrid methods directly applied to the discretized problem as an outer iteration based on appropriate smoothers (as a sort of inner iteration). For PDE-constrained optimization problems this approach is also known as one-shot multigrid strategy, see Ta'asan [105]. One of the most important ingredients of such a multigrid method is an appropriate smoother.

So far, the multigrid convergence analysis for KKT systems of PDE-constrained optimization problems is not as developed as for elliptic PDEs. One line of argument exploits the fact that the KKT system is a compact perturbation of an elliptic problem. This guarantees the convergence of the multigrid method if the coarse grid is sufficiently fine, see Borzi, Kunisch and Kwak [18]. A second strategy is based on a Fourier analysis, which, strictly speaking, covers only the case of uniform meshes with special boundary conditions (and small perturbations of this situation), see, e.g., Borzi, Kunisch and Kwak [18], Arian and Ta'asan [2].

A typical elliptic optimal control looks like the following: Let Ω be a bounded convex polygonal domain in $\mathbb{R}^2$. Let $L^2(\Omega)$ and $H^1(\Omega)$ denote the usual Lebesgue space and Sobolev space, respectively. The goal is to find the state $y \in H^1(\Omega)$ and the control $u \in L^2(\Omega)$ such that

$$J(y,u) = \min_{(z,v)\in H^1(\Omega)\times L^2(\Omega)} J(z,v)$$

with cost functional

$$J(z,v) = \frac{1}{2}\|z - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|v\|^2_{L^2(\Omega)}$$

subject to the state equations

$$\begin{aligned} -\Delta y + y &= u \qquad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 \qquad \text{on } \Gamma, \end{aligned}$$

where Γ denotes the boundary of Ω, $y_d \in L^2(\Omega)$ is the desired state and $\nu > 0$ is the weight of the cost of the control.

By introducing the adjoint state $p \in H^1(\Omega)$ we get the optimality system, see, e.g., Tröltzsch [107] and the weak formulation leads to the following mixed variational problem: Find $x = (y, u) \in X = Y \times U$ with $Y = H^1(\Omega)$, $U = L^2(\Omega)$ and $p \in Q = H^1(\Omega)$ such that

$$\begin{array}{lcll} a(x,w) & + & b(w,p) = \langle F, w\rangle & \text{for all } w \in X, \\ b(x,q) & & = 0 & \text{for all } q \in Q \end{array}$$

with

$$\begin{aligned} a(x,w) &= (y,z)_{L^2(\Omega)} + \nu(u,v)_{L^2(\Omega)}, \\ b(w,q) &= (z,q)_{H^1(\Omega)} - (v,q)_{L^2(\Omega)}, \\ \langle F, w\rangle &= (y_d, z)_{L^2(\Omega)}, \end{aligned}$$

where $w = (z, v)$ with $z \in Y$, $v \in U$, and $(\cdot,\cdot)_H$ is the standard scalar product in a Hilbert space H, whose norm is denoted by $\|\cdot\|_H$.

Let $(\mathcal{T}_k)$ be a sequence of triangulations of Ω. We consider the following discretization by continuous and piecewise linear finite elements:

$$\begin{aligned} X_k = Y_k \times U_k &= \{(z,v) \in C(\bar\Omega) \times C(\bar\Omega) : z|_T, v|_T \in \mathcal{P}_1 \ \text{ for all } T \in \mathcal{T}_k\}, \\ Q_k &= \{q \in C(\bar\Omega) : q|_T \in \mathcal{P}_1 \ \text{ for all } T \in \mathcal{T}_k\}, \end{aligned}$$

where $\mathcal{P}_1$ denotes the polynomials of total degree less or equal to 1. Then we obtain the following discrete variational problem: Find $x_k \in X_K$ and $p_k \in Q_k$ such that

$$\begin{array}{lcll} a(x_k,w_k) & + & b(w_k,p_k) = \langle F, w_k\rangle & \text{for all } w_k \in X_k, \\ b(x_k,q_k) & & = 0 & \text{for all } q_k \in Q_k \end{array}$$

By introducing the standard nodal basis, we finally obtain the following saddle point problem in matrix-vector notation:

$$\mathcal{K}_k \begin{pmatrix} \underline{x}_k \\ \underline{p}_k \end{pmatrix} = \begin{pmatrix} \underline{f}_k \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_k = \begin{pmatrix} A_k & B_k^T \\ B_k & 0 \end{pmatrix}$$

where

$$A_k = \begin{pmatrix} M_k & 0 \\ 0 & \nu M_k \end{pmatrix} \quad \text{and} \quad B_k = \begin{pmatrix} K_k & -M_k \end{pmatrix}.$$

Here M_k denotes the mass matrix representing the $L^2(\Omega)$ scalar product on Y_k and K_k denotes the stiffness matrix representing the $H^1(\Omega)$ scalar product on Y_k.

Schwarz-type Smoother

One of the most important ingredients of a multigrid method is an appropriate smoother. A first approach for constructing such smoothers is to combine standard smoothers applied to the components elliptic state and adjoint equations complemented with a special relaxation method for the control equation, see, e.g., Arian and Ta'asan [2].

A second class of smoothers are point smoothers, where the variables are grouped pointwise (with respect to the nodes of the underlying mesh) and one or several sweeps of point-block Jacobi or point-block Gauß-Seidel sweeps with respect to this grouping are performed, see, e.g., Borzi, Kunisch and Kwak [18].

A natural extension of point smoothers are patch smoothers: The computational domain is divided into small (overlapping or non-overlapping) patches. One iteration step of the smoothing process consists of solving local mixed problems on each patch one-by-one either in a Jacobi-type or Gauß-Seidel-type manner. This results in an additive or multiplicative Schwarz-type smoother. The technique was successfully used for the Navier-Stokes equations, see Vanka [108]. The general construction and the analysis of patch smoothers for mixed problems was discussed in Schöberl and Zulehner [85], where a particular patch smoother was proposed for the Stokes problem. A straight forward application of this construction to KKT systems for elliptic control problems fails, since an essential feature exploited in the multigrid convergence analysis of the Stokes problem was the positivity of (1,1)-block everywhere, whereas in optimal control problems the (1,1)-block is usually positive only on the kernel of the (2,1)-block.

We describe here only the main idea of constructing such a Schwarz-type patch smoother for elliptic optimal control problems. For more details we refer to Simon and Zulehner [92]. Since the smoothing procedure involves only one level k of the hierarchy of spaces, we will simplify the notation by dropping the subscript k and omitting underlining the vectors. So we discuss iterative methods (as smoothers) for linear systems of equations of the form:

$$\mathcal{K}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{with} \quad \mathcal{K} = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}, \tag{16}$$

where $x \in \mathbb{R}^n, p \in \mathbb{R}^m$.

For setting up local sub-problems a set of linear operator is introduced:

$$P_i : \mathbb{R}^{n_i} \to \mathbb{R}^n, \quad Q_i : \mathbb{R}^{m_i} \to \mathbb{R}^m, \quad \text{for } i = 1, \ldots, N,$$

where the dimensions n_i and m_i are typically much smaller than the dimensions n and m of the original spaces, respectively. Under some proper conditions we have complete space decompositions

$$\mathbb{R}^n = \sum_{i=1}^{N} P_i(\mathbb{R}^{n_i}) \quad \text{and} \quad \mathbb{R}^m = \sum_{i=1}^{N} Q_i(\mathbb{R}^{m_i}),$$

and, additionally, the prolongations P_i determine a special partition of unity.

For each index $i = 1, \ldots, N$, local matrices $\hat{A}_i$, B_i and $\hat{S}_i$ have to be chosen, which determine local matrices $\hat{\mathcal{K}}_i$ of the form

$$\hat{\mathcal{K}}_i = \begin{pmatrix} \hat{A}_i & B_i^T \\ B_i & B_i \hat{A}_i^{-1} B_i^T - \hat{S}_i \end{pmatrix}.$$

With the help of the local saddle point matrices $\hat{\mathcal{K}}_i$ the following iterative method is constructed: Starting from some approximations $x^{(j)}$ and $p^{(j)}$ of the exact solutions x and p of (16) we consider iterative methods of the form:

$$x^{(j+1)} = x^{(j)} + \sum_{i=1}^{N} P_i s_i^{(j)}, \quad p^{(j+1)} = p^{(j)} + \sum_{i=1}^{N} Q_i r_i^{(j)},$$

where $(s_i^{(j)}, r_i^{(j)}) \in \mathbb{R}^{n_i} \times \mathbb{R}^{m_i}$ solves a local saddle point problem of the form

$$\hat{\mathcal{K}}_i \begin{pmatrix} s_i^{(j)} \\ r_i^{(j)} \end{pmatrix} = \begin{pmatrix} P_i^T[f - Ax^{(j)} - B^T p^{(j)}] \\ Q_i^T[g - Bx^{(j)} + Cp^{(j)}] \end{pmatrix} \quad \text{for all } i = 1, \ldots, N.$$

The crucial point for the analysis of the smoothing property is, that the local matrices are related to the global matrices via some proper commutativity conditions. This iterative method can then be written equivalently as the following preconditioned Richardson method:

$$x^{(j+1)} = x^{(j)} + s^{(j)}, \quad p^{(j+1)} = p^{(j)} + r^{(j)},$$

where $(s^{(j)}, r^{(j)})$ solve the equation

$$\hat{\mathcal{K}} \begin{pmatrix} s^{(j)} \\ r^{(j)} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} - \mathcal{K} \begin{pmatrix} x^{(j)} \\ p^{(j)} \end{pmatrix} \quad \text{with} \quad \hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix}. \tag{17}$$

So the additive Schwarz-type iterative method can be represented as an symmetric inexact Uzawa method. Let $\mathcal{M}$ denote the associated iteration matrix, given by

$$\mathcal{M} = I - \hat{\mathcal{K}}^{-1}\mathcal{K},$$

which controls the error propagation for the iterative method.

Multigrid convergence analysis

A classical technique for analyzing the convergence of multigrid methods relies on two properties: the approximation property and the smoothing property, see Hackbusch [37]. The approximation property measures the effect of the coarse grid correction, whereas the smoothing property measures the effect of the smoothing procedure, respectively, usually measured in properly chosen discrete norms. The smoothing property can be translated to the following condition in matrix-notation:

$$\|\mathcal{K}\mathcal{M}^k\|_{\mathcal{L}} \leq \eta(m)\|\mathcal{K}\|_{\mathcal{L}},$$

for some function $\eta(m)$ which is independent of the level k, and

$$\eta(m) \to 0 \quad \text{for } m \to \infty,$$

and $\mathcal{L}$ is a symmetric and positive definite matrix which represents the mesh-dependent (L^2-like) norm.

It was shown in Schöberl and Zulehner [85] that, given a symmetric and positive definite matrix $\hat{A}$ and $\hat{S}$, satisfying

$$\hat{A} \geq A \quad \text{and} \quad \hat{S} \geq C + B\hat{A}^{-1}B^T,$$

the following estimate holds:

$$\|\mathcal{K}\mathcal{M}^m\|_{\mathcal{L}} \leq \eta_0(m)\|\mathcal{D}\|_{\mathcal{L}},$$

where $\mathcal{K}$ is given by (16), $\hat{\mathcal{K}}$ is given by (17), $\mathcal{D}$ is given by

$$\mathcal{D} = \begin{pmatrix} \hat{A} - A & 0 \\ 0 & \hat{S} - C - B\hat{A}^{-1}B^T \end{pmatrix},$$

$\mathcal{L}$ is an arbitrary symmetric and positive definite matrix, and

$$\eta_0(m) = \frac{1}{2^{m-1}}\binom{m-1}{[m]/2} \leq \begin{cases} \sqrt{\frac{2}{\pi(m-1)}} & \text{for even } m \\ \sqrt{\frac{2}{\pi m}} & \text{for odd } m. \end{cases}$$

Here $\binom{n}{k}$ denotes the binomial coefficient and $[x]$ denotes the largest integer smaller than or equal to $x \in \mathbb{R}$.

That means, the smoothing property is satisfied for the additive Schwarz-type method, if the local problems are constructed in such a way that the associated global matrices $\hat{A}$,B and $\hat{S}$ satisfy the conditions

$$\hat{A} \geq A \quad \text{and} \quad \hat{S} \geq C + B\hat{A}^{-1}B^T \tag{18}$$

and if, additionally, the following scaling condition holds:

$$\|\mathcal{D}\|_{\mathcal{L}} \leq c_R\|\mathcal{K}\|_{\mathcal{L}} \quad \text{with} \quad \mathcal{D} = \begin{pmatrix} \hat{A} - A & 0 \\ 0 & \hat{S} - C - B\hat{A}^{-1}B^T \end{pmatrix} \tag{19}$$

for some constant c_R independent of the level k.

In Schöberl and Zulehner [85] this strategy was successfully applied to the Stokes problem, discretized by the Crouzeix-Raviart mixed finite element method. For the global matrix $\hat{A}$ a constant multiple of $\operatorname{diag} A$ was chosen. Special local matrices were constructed and all requirements of the analysis could be verified. In particular, the scaling condition (19) could be shown. For a typical class of problems from optimal control, like the one presented above, the same choice of $\hat{A}$ leads to a violation of the scaling condition (19). We will now show how the construction must be modified to keep the right scaling without losing any of the other requirements.

Let N be the number of nodes of the triangulation $\mathcal{T}$. For each $i = 1, \ldots, N$ representing a node of the triangulation, let $\mathcal{N}_i$ be the set of all indices consisting of i and the indices of all neighboring nodes. Then the associated local patch consists of all unknowns

y and u which are associated to node with indices from $\mathcal{N}_i$ and the unknown of p which is associated to the node with index i, see Figure 14 for an illustration of a local patch. The corresponding prolongations are the canonical embeddings into $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Additionally, an appropriate scaling is required which takes into account the local overlap depth of the components of the primal variables.

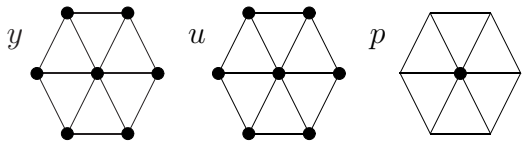

Figure 14: local patch

Next we have to choose a matrix $\hat{A}$. It seems to be natural to choose

$$\hat{A} = \frac{1}{\sigma} \operatorname{diag} A = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} M & 0 \\ 0 & \nu \operatorname{diag} M \end{pmatrix}$$

with a suitable parameter $\sigma > 0$. But in order to prove the smoothing property, we have to check, if the estimate (19) is fulfilled. This is not the case with this definition of $\hat{A}$ for parameters $\sigma = \mathcal{O}(1)$. Instead we choose

$$\hat{A} = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} K & 0 \\ 0 & \nu \operatorname{diag} M \end{pmatrix}$$

with σ small enough to ensure

$$\hat{A} \geq A.$$

The local matrices are then constructed by restricting the global matrices to the local patches.

By determining the structure of the patches and the local problems, the smoothing procedure is defined. For this additive Schwarz smoother, the smoothing property holds with a smoothing rate $\eta(m) = \mathcal{O}(1/\sqrt{m})$. In Simon and Zulehner [92] we were also able to prove the approximation property, which finally led to a rigorous convergence analysis of the corresponding multigrid method.

We tested the multigrid method on the unit square and homogeneous data $y_d = 0$. Randomly chosen starting values were used. The discretized problem was solved by a multigrid iteration with a W-cycle and $m/2$ pre- and $m/2$ post-smoothing steps. The multigrid iteration was performed until the Euclidean norm of the solution was reduced by a factor $\epsilon = 10^{-8}$. Table 5 contains the total number of unknowns, the number of iterations and the (average) convergence rates depending on the level and the number of smoothing steps. It shows a typical multigrid convergence behavior, namely the independence of the grid level and the expected improvement of the rates with an increasing number of smoothing steps.

		smoothing steps							
level	$n_h + m_h$	5+5		7+7		10+10		15+15	
5	3 267	46	0.668	30	0.538	21	0.411	17	0.330
6	12 675	48	0.679	34	0.578	24	0.455	17	0.330
7	49 923	49	0.685	35	0.587	25	0.467	17	0.333
8	198 147	49	0.685	35	0.588	25	0.469	17	0.333
9	789 507	49	0.685	35	0.589	25	0.469	17	0.334

Table 5: Convergence rates for the additive Schwarz smoother

		smoothing steps							
level	$n_h + m_h$	5+5		7+7		10+10		15+15	
5	3 267	21	0.410	19	0.372	16	0.307	12	0.208
6	12 675	22	0.420	20	0.382	16	0.310	12	0.214
7	49 923	22	0.423	19	0.378	16	0.311	12	0.213
8	198 147	22	0.423	20	0.383	16	0.313	12	0.213
9	789 507	22	0.423	20	0.383	16	0.314	12	0.214

Table 6: Convergence rates for the multiplicative Schwarz smoother

Table 6 shows the convergence rates with the multiplicative version of the smoother. As expected, the rates are significantly better than the rates for the additive smoother. The number of smoothing steps which are necessary to achieve convergence on all levels is much smaller than in the additive version. However, a theoretical analysis for the convergence and smoothing properties is still missing.

In a more recent work we considered the reduced KKT system by eliminating the control u. The resulting system can be seen as a positive definite, but nonsymmetric system (already in matrix-vector notation):

$$\begin{pmatrix} \nu K & -M \\ M & K \end{pmatrix} \begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}.$$

Actually, this is the system, which was solved in Borzi, Kunisch and Kwak [18] using the point smoother for the multigrid method. But it can be equivalently formulated as a symmetric but indefinite system:

$$\begin{pmatrix} \nu K & -M \\ -M & -K \end{pmatrix} \begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f \\ -g \end{pmatrix},$$

which fits perfectly in the class of symmetric saddle point problems, considered in Schöberl and Zulehner [85]. Here we do not need the augmentation technique, presented above, since the upper left block is coercive on the whole space. Setting up the

proper local problems, we can again write the whole smoothing procedure as a preconditioned Richardson method in a global way. Furthermore, we introduce an additional overrelaxation parameter ω, leading to the following iteration:

$$\begin{pmatrix} y^{(j+1)} \\ p^{(j+1)} \end{pmatrix} = \begin{pmatrix} y^{(j)} \\ p^{(j)} \end{pmatrix} + \omega \hat{\mathcal{K}}^{-1} \left[\begin{pmatrix} f \\ -g \end{pmatrix} - \mathcal{K} \begin{pmatrix} y^{(j)} \\ p^{(j)} \end{pmatrix} \right].$$

Using a convergence result for the relaxed method together with an extension of Reusken's Lemma, see Ecker and Zulehner [31], it can be shown that for all relaxation factors $\omega \in (0, 2)$ the smoothing property is fulfilled. As a matter of fact numerical experiments show a better performance if we choose a relaxation parameter $\omega > 1$, e.g., $\omega = 1.6$, see Table 7 and Table 8.

		smoothing steps							
level	$n_h + m_h$	1+1		2+2		3+3		5+5	
5	2 178	27	0.502	14	0.267	10	0.156	7	0.068
6	8 450	27	0.502	14	0.264	10	0.156	7	0.068
7	33 282	27	0.502	14	0.265	11	0.162	7	0.069
8	132 098	27	0.502	14	0.265	10	0.158	7	0.069
9	526 338	27	0.502	14	0.265	11	0.163	7	0.069

Table 7: Convergence rates using relaxation parameter $\omega = 1$

		smoothing steps							
level	$n_h + m_h$	1+1		2+2		3+3		5+5	
5	2 178	16	0.301	9	0.127	7	0.067	5	0.023
6	8 450	16	0.302	9	0.128	7	0.066	5	0.024
7	33 282	16	0.302	10	0.135	7	0.067	5	0.024
8	132 098	16	0.302	10	0.135	7	0.067	5	0.024
9	526 338	16	0.302	10	0.135	7	0.068	5	0.024

Table 8: Convergence rates using relaxation parameter $\omega = 1.6$

Symmetric Indefinite Preconditioners for KKT Systems

In Schöberl and Zulehner [87] we contributed to the first approach and considered large scale sparse linear systems of equations in saddle point form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{20}$$

where A is a real, symmetric and positive semi-definite n-by-n matrix with

$$(Aw, w) > 0 \quad \text{for all } w \in \ker B \text{ with } w \neq 0, \tag{21}$$

and (x, w) denotes the Euclidean scalar product, B is a real m-by-n matrix with full rank $m \leq n$, and B^T denotes the transposed matrix of B. These conditions guarantee that the matrix

$$\mathcal{K} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$$

is non-singular. A well-known class of preconditioners is given by

$$\hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix},$$

where $\hat{A}$ and $\hat{S}$ are symmetric and positive definite matrices, see Bank, Welfert and Yserentant [4]. Estimates for the extreme eigenvalues of $\hat{\mathcal{K}}^{-1}\mathcal{K}$ were derived in Zulehner [114] under the assumption that A is positive definite on the whole space. However, the estimate for the smallest eigenvalue degenerates, if directly applied to the case considered here. This gap is closed in this paper:

Theorem 1 *Assume that $A \geq 0$, condition (21) is satisfied, and* $\operatorname{rank} B = m$. *Let $\hat{A} > 0$ and $\hat{S} > 0$ with*

$$(Aw, w) \geq \alpha\,(\hat{A}w, w) \quad \textit{for all } w \in \ker B \quad \textit{and} \quad \hat{A} \geq A, \tag{22}$$

and

$$\hat{S} \leq B\hat{A}^{-1}B^T \leq \beta\,\hat{S} \tag{23}$$

with constants α and β with $0 < \alpha \leq 1$ and $0 < \beta \leq 1$. Then

$$\lambda_{max}(\hat{\mathcal{K}}^{-1}\mathcal{K}) \leq \beta + \sqrt{\beta^2 - \beta} = \beta\,(1 + \sqrt{1 - 1/\beta})$$

and

$$\begin{aligned} \lambda_{min}(\hat{\mathcal{K}}^{-1}\mathcal{K}) &\geq \frac{1}{2}\left[2 + \alpha - 1/\beta - \sqrt{(2 + \alpha - 1/\beta)^2 - 4\alpha}\right] \\ &\geq \alpha\left[\frac{2}{\sqrt{1 - 1/\beta} + \sqrt{5 - 1/\beta}}\right]^2 > 0. \end{aligned}$$

By slightly strengthening the conditions (22) and (23) to

$$(Aw, w) \geq \alpha\,(\hat{A}w, w) \quad \text{for all } w \in \ker B \quad \text{and} \quad \hat{A} > A \tag{24}$$

and

$$\hat{S} < B\hat{A}^{-1}B^T \leq \beta\,\hat{S}, \tag{25}$$

the scalar product

$$\left(\begin{pmatrix} x \\ p \end{pmatrix}, \begin{pmatrix} w \\ q \end{pmatrix}\right)_{\mathcal{D}} = ((\hat{A} - A)x, w) + ((B\hat{A}^{-1}B^T - \hat{S})p, q). \tag{26}$$

is well-defined, and the standard conjugate gradient method can be applied to the preconditioned system

$$\hat{\mathcal{K}}^{-1}\mathcal{K}\begin{pmatrix} x \\ p \end{pmatrix} = \hat{\mathcal{K}}^{-1}\begin{pmatrix} f \\ g \end{pmatrix} \tag{27}$$

with respect to the scalar product (26). It is well-known, that the error $e^{(k)}$ for the k-th iterate $(x^{(k)}, p^{(k)})^T$ measured in the corresponding energy norm can be estimated by

$$e^{(k)} \leq \frac{2q^k}{1+q^{2k}}\, e^{(0)} \quad \text{with} \quad q = \frac{\sqrt{\kappa(\hat{\mathcal{K}}^{-1}\mathcal{K})} - 1}{\sqrt{\kappa(\hat{\mathcal{K}}^{-1}\mathcal{K})} + 1},$$

where $\kappa(\hat{\mathcal{K}}^{-1}\mathcal{K})$ denotes the relative condition number:

$$\kappa(\hat{\mathcal{K}}^{-1}\mathcal{K}) = \frac{\lambda_{\max}(\hat{\mathcal{K}}^{-1}\mathcal{K})}{\lambda_{\min}(\hat{\mathcal{K}}^{-1}\mathcal{K})}.$$

From Theorem 1 the following upper bound for the relative condition number follows:

$$\begin{aligned} \kappa(\hat{\mathcal{K}}^{-1}\mathcal{K}) &\leq \frac{2(\beta + \sqrt{\beta^2 - \beta})}{2 + \alpha - 1/\beta - \sqrt{(2+\alpha-1/\beta)^2 - 4\alpha}} \equiv \kappa(\alpha, \beta) \\ &\leq \frac{\beta}{\alpha}\,(1 + \sqrt{1 - 1/\beta}) \left[\frac{\sqrt{1-1/\beta} + \sqrt{5 - 1/\beta}}{2}\right]^2. \end{aligned}$$

This shows that the convergence rate q can be bounded by α and β only. If the preconditioners are chosen such that α and β are independent of certain parameters like the mesh size h of some discretization or some involved regularization parameter ν, then the convergence rate is also robust with respect to such parameters.

We applied the general result to the following optimization problem with PDE-constraints, where $\Omega \subset \mathbb{R}^d$ is an open and bounded set:
Find the state $y \in H^1(\Omega)$ and the control $u \in L^2(\Omega)$ such that

$$J(y, u) = \min_{(z,v)\in H^1(\Omega)\times L^2(\Omega)} J(z, v),$$

subject to the state equation with distributed control u

$$\begin{aligned} -\Delta y + y &= u \quad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where the cost functional is given by

$$J(y,u) = \frac{1}{2}\int_\Omega (y-y_d)^2\,dx + \frac{\nu}{2}\int_\Omega u^2\,dx.$$

By discretising the associated KKT system by a standard finite element method, we finally obtain the following saddle point problem in matrix-vector notation:

$$\begin{aligned} A_h\underline{x}_h + B_h^T\underline{p}_h &= \underline{f}_h,\\ B_h\underline{x}_h &= 0, \end{aligned}$$

with

$$A_h = \begin{pmatrix} M_h & 0\\ 0 & \nu\, M_h\end{pmatrix} \quad\text{and}\quad B_h = \begin{pmatrix} K_h & -M_h\end{pmatrix},$$

where M_h denotes the mass matrix representing the $L^2(\Omega)$ inner product on Y_h, and K_h denotes the stiffness matrix representing the bilinear form (on Y) of the state equation, here $(\nabla y,\nabla q)_{L^2(\Omega)} + (y,q)_{L^2(\Omega)}$, on Y_h.

The suggested preconditioner

$$\hat{\mathcal{K}}_h = \begin{pmatrix} \hat{A}_h & B_h^T\\ B_h & B_h\hat{A}_h^{-1}B_h^T - \hat{S}_h\end{pmatrix} \quad\text{for}\quad \mathcal{K}_h = \begin{pmatrix} A_h & B_h^T\\ B_h & 0\end{pmatrix}$$

is given by

$$\hat{A}_h = \frac{1}{\sigma}\,\hat{X}_h = \frac{1}{\sigma}\begin{pmatrix}\hat{Y}_h & 0\\ 0 & \nu\,\hat{M}_h\end{pmatrix} \quad\text{and}\quad \hat{S}_h = \frac{\sigma}{\tau}\,\frac{1}{\nu}\,\hat{Y}_h \tag{28}$$

with real parameters $\sigma > 0$ and $\tau > 0$, where $\hat{Y}_h$ is a preconditioner for the stiffness matrix $Y_h = \sqrt{\nu}\,K_h + M_h$ of the bilinear form $\sqrt{\nu}\,(\nabla y,\nabla q)_{L^2(\Omega)} + (\sqrt{\nu}+1)\,(y,q)_{L^2(\Omega)}$ and a simple preconditioner $\hat{M}_h$ for the mass matrix M_h.

It is reasonable to assume that

$$(1-q_X)\,\hat{Y}_h \le \underline{Y}_h \le \hat{Y}_h \quad\text{and}\quad (1-q_X)\,\hat{M}_h \le M_h \le \hat{M}_h,$$

for some small value $q_X \in [0,1)$. The factor q_X describes the quality of the preconditioners $\hat{Y}_h$ and $\hat{M}_h$.

It can be shown that the conditions (24) and (25) are satisfied with

$$\alpha = \sigma\,(1-q_X)\,\frac{2}{3} \quad\text{and}\quad \beta = \tau$$

for parameters σ and τ satisfying

$$\sigma < 1 \quad\text{and}\quad \tau > \frac{4}{3(1-q_X)^2}.$$

In particular, assuming that $q_X \approx 0$, we can except $\alpha \approx 2/3$ and $\beta \approx 4/3$ for $\sigma \approx 1$ and $\tau \approx 4/3$, leading to a rough estimate of the condition number $\kappa \approx \kappa(2/3, 4/3) \approx 4$, which implies a convergence factor $q \approx 1/3$ for the conjugate gradient method.

This was confirmed by numerical experiments for $\Omega = (0,1)^3$, where $\hat{Y}_h$ is one V-cycle of the multigrid method with m_1 forward Gauss-Seidel steps for pre-smoothing and m_1 backward Gauss-Seidel steps for post-smoothing (in short $V(m_1, m_1)$) for the second-order elliptic differential operator represented by the bilinear form $\sqrt{\nu}\,(\nabla y, \nabla q)_{L^2(\Omega)} + (\sqrt{\nu} + 1)\,(y, q)_{L^2(\Omega)}$. For $\hat{M}_h$ we use m_2 steps of the symmetric Gauss-Seidel method (in short $SGS(m_2)$). Table 9 shows that the number of iterations does not depend on the level of refinement. L denotes the level of refinement, $n + m$ the total number of all unknowns $\underline{y}_h$, $\underline{u}_h$ and $\underline{p}_h$, k the number of iterations needed to satisfy the stopping rule

$$r^{(k)} \leq \varepsilon\, r^{(0)} \quad \text{with } \varepsilon = 10^{-8},$$

and t the total cpu time in seconds.

Table 9: Dependence of the number of iterations on the mesh size for fixed $\nu = 1$.

level L	number of unknowns $n + m$	iterations k	cpu time t (in seconds)
3	1,107	14	0.06
4	7,395	15	0.61
5	53,955	15	6.96
6	412,035	16	62.04
7	3,200,227	15	559.16

Table 10 shows that the number of iterations does not depend on the regularization parameter ν either. The results are given for refinement level $L = 5$.

Table 10: Dependence of the number of iterations on ν for fixed refinement level $L = 5$.

ν	iterations k
10^{-4}	15
10^{-2}	14
1	15
10^{2}	14
10^{4}	15

1.4 Collaboration Within and Outside the SFB

1.4.1 Cooperations Inside the SFB / Internal Cooperations

- **1304:** D. Lukas (F1309) and N. Bila (F1304) were looking for a use of symmetries for a dimensional reduction of the magnetostatic state problem. However,

the reduction is based on Fourier analysis and only a constant material is allowed. Moreover, only geometrical, e.g. cylindrical, symmetries were proposed to be explored, which would surely lead to analytical solutions well-known from physics.

- **1308:** A very successful cooperation was established with Subproject F1308. Here we devoloped a new approach to topology optimization, with local stess constraints, one of the hottest topics in this field. We successfully combined an all-at-once formulation, reformulation of constraints and a phase–field regularization to a new solution method. This approach is described in more detail in the above Subsection 1.3.1 and resulted in a joint paper Burger and Stainko [26].

  D. Lukas (F1309) and M. Burger (F1308) tried to employ a phase-field method for an interface shape optimization. However, both collaborators independently shared the numerical evidence of an extremely slow convergence. Moreover, one has to tune sensitively the value of the phase-field penalty with respect to the penalization of inequality constraints. Therefore, this joint work has stalled.

- **1315:** D. Lukas (F1309) with M. Shallaby (F1315) attempted to use an implicit B-spline geometry representation contructed in a wavelet hierarchy for multilevel interface shape optimization. This approach lead to a higher-orderpolynomial representation of the material function. Howver, due to a very technical implementation this approach was not realized, and we rather focused on a hierarchical parameteric representation of shapes, see the joint paper Lukas and Chalmoviansky [55].

  Furthermore, Pechstein and Jüttler [65] proposed new monotonicity-preserving interproximation of B-H – curves that is used in our Maxwell-code for solving non-linear magnetic field problems. The corresponding software product was bought – and is currently used – by the Robert Bosch GmbH Stuttgart.

1.4.2 Cooperations Outside the SFB / External Cooperations

- **Prof. Dr. Zdeněk Dostál** (VSB-Technical University of Ostrava, Czech Republic):
  Z. Dostal is a professor of applied mathematics and the head of the Department of Applied Mathematics at TU Ostrava. We cooperate on development of optimal complexity augmented Lagrangian algorithms for equality and box constrained quadratic programming with the aim to efficiently solve KKT systems arising in optimal control and topology optimization.

  In augmented Lagrangians the augmented penalty (regularization) parameter is traditionally increased, since the regularized solutions are known to converge to the original one. However, this makes the resulting augmented operators less well-posed. In the approach of Dostál [29] there are two important differences. Firstly, the augmented parameter is increased only in cases when the Lagrange functional does not grow more than of a given amount (semi-monotonicity), which is due to a

careful analysis. Secondly, we solve the inner linear systems only up to a precision proportional to the violence of the equality constraint. Then, one can prove a well–controlled bound of the augmented Lagrange parameter which depends only on the smallest eigenvalue of the Hessian. This implies a uniform equivalence of all the Hessians of the augmented quadratics. A similar well-controlled estimate is given to the number of outer iterations. Thus, in a combination with a multigrid preconditioning to the primal and dual inner products as well as to the Hessian, we get the linear complexity, provided the linear complexity of the matrix-vector products.

- **Dr. J. Krauss** (RICAM):
  J. Kraus is a specialist on algebraic multigrid (AMG) solution techniques. We cooperate on applications of AMG as a state problem solver in optimal interface shape design. Traditionally in the nested approach, design perturbations causes perturbations of grid and, consequently, of the state stiffness matrix, which has to be re-preconditioned. However, this approach can be performed unless the grid deformation is too bad. Another approach is to keep the grid and allow jumping material coefficients. From our first results it turns out that AMG preconditioning is an efficient technique for the latter approach.

- **Prof. Dr. M.P. Bendsøe** (Technical University of Denmark):
  R. Stainko continued the scientific cooperation with Prof. Bendsøe, which always led to fruitful discussions about the topic of topology optimization. Moreover, Prof. Bendsøe invited R. Stainko to DTU to give a talk in their TOPOPT seminars about the approach to local stress constraints.

- **Prof. Dr. O. Sigmund** (Technical University of Denmark):
  We also continued our cooperation with Prof. Sigmund. Together with Prof. Bendsøe and other researchers he forms the so called TOPOPT group at the DTU, where in regular meetings a lot of knowledge concerning applied topology optimization and theory is exchanged. The fruitful scientific relation resulted in a post-doc position for R. Stainko at the DTU, invited by Prof. Sigmund.

- **Dr. M. Stolpe** (Technical University of Denmark):
  During his stays at DTU, R. Stainko started a cooperation with Dr. Stolpe, whose idea of constraint reformulation is one of the cornerstones of the new approach of Dr. Burger and R. Stainko. M. Stolpe combines a huge knowledge about optimization theory and structural optimization and is an invalueable contact partner for hot and up-to-date topics like stress constrained optimization.

- **Prof. Dr. A. Rösch** (University Duisburg-Essen, Germany): R. Simon cooperated with him on error estimates and superconvergence properties for optimal control problems, Rösch and Simon [76, 77].

References

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

[2] E. Arian and S. Ta'asan. Multigrid one-shot methods for optimal control problems: Infinite dimensional control. ICASE-Report 94-52, NASA Langley Research Center, Hampton VA, 1994.

[3] N. V. Banichuk. *Introduction to Optimization of Structures*. Springer, New York, 1990.

[4] R. Bank, B. Welfert, and H. Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56:645–666, 1990.

[5] P. D. Barba, P. Navarra, A. Savini, and R. Sikora. Optimum design of iron–core electromagnets. *IEEE Trans. Magnetics*, 26:646–649, 1990.

[6] A. Battermann and M. Heinkenschloss. Preconditioners for Karush-Kuhn-Tucker matrices arising in the optimal control of distributed systems. In W. Desch et al., editors, *Control and Estimation of Distributed Parameter Systems*, Int. Ser. Numer. Math. 126, pages 15–32. Birkhäuser, Basel, 1998.

[7] A. Battermann and E. Sachs. Block preconditioners for KKT systems in PDE-governed optimal control problems. In K.-H. Hoffmann et al., editors, *Fast Solution of Discretized Optimization Problems*, Int. Ser. Numer. Math. 138, pages 1–18. Birkhäuser, Basel, 2001.

[8] D. Begis and R. Glowinski. Application de la méthode des éléments finis à la résolution d'un probleème de domaine optimal. *Appl. Math. Opt.*, 2:130–169, 1975.

[9] M. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications*. Springer, Berlin, 2003.

[10] M. P. Bendsøe. Optimal shape design as a material distribution problem. *Structural Optimization*, 1:193–202, 1989.

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

[12] M. P. Bendsøe and O. Sigmund. Material interpolation schemes in topology optimization. *Archive of Applied Machincs*, 69:635–654, 1999.

[13] M. P. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications*. Springer, Berlin, 2003.

[14] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver. *SIAM J.Sci.Comput.*, 27(2):687–713, 2005.

[15] A. D. Börner. *Mathematische Modelle für Optimum Shape Design Probleme in der Magnetostatik*. PhD thesis, Universität Hamburg, Germany, 1985.

[16] T. Borrvall. *Computational Topology Optimization of Elastic Continua by Design Restrictions*. PhD thesis, Linköping University, Sweden, 2000. Theses No. 848.

[17] T. Borrvall and J. Petersson. Topology optimization using regularized intermediate density control. *Computer Methods in Applied Mechanics and Engineering*, 190:4911–4928, 2001.

[18] A. Borzi, K. Kunisch, and D. Kwak. Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system. *SIAM J. Control Optimization*, 41(5):1477–1497, 2003.

[19] B. Bourdin. Filters in topology optimization. *International Journal for Numerical Methods in Engineering.*, 50:2134–2158, 2000.

[20] R. B. Brandtstätter, W. Ring, C. Magele, and K. R. Richter. Shape design with great geometrical deformations using continuously moving finite element nodes. *IEEE Transactions on Magnetics*, 34(5):2877–2880, September 1998.

[21] M. Bruyneel, P. Duysinx, and C. Fleury. A family of MMA approximations for structural optimization. *Structural and Multidisciplinary Optimization*, 24:263–276, 2002.

[22] D. Bucur and J.-P. Zolesio. Existence and stability of the optimum in shape optimization. *ZAMM*, 76(2):77–80, 1996.

[23] M. Burger and W. Mühlhuber. Iterative regularization of parameter identification problems by SQP methods. *Inverse Problems*, 18:943 – 970, 2002.

[24] M. Burger and W. Mühlhuber. Numerical approximation of a SQP-type method for parameter identification. *SIAM J. Numer. Anal.*, 40(5):1775 – 1797, 2002.

[25] M. Burger and R. Stainko. Phase-field relaxation of topology optimization with local stress constraints. Technical Report 2004-35, J. Kepler University Linz, SFB F013, 2004.

[26] M. Burger and R. Stainko. Phase–field relaxation of topology optimization with local stress constraints. *SIAM Journal on Control and Optimization*, 45(4):1447–1466, 2006.

[27] A. V. Cherkaev. *Variational Methods for Structural Optimization*. Springer, New York, Berlin, Heidelberg, 2000.

[28] M. C. Delfour and J.-P. Zolesio. *Shape and Geometries: Analysis, Differential Calculus, and Optimization*, volume 4 of *Advances in Design and Control*. SIAM, 2001.

[29] Z. Dostál. Semi-monotonic inexact augmented Lagrangians for quadratic programming with equality constraints. *Optimization Methods and Software*, 20:715–727, 2005.

[30] T. Dreyer, B. Maar, and V. Schulz. Multigrid optimization in applications. *J. Comput. Appl. Math.*, 120(1-2):67–84, 2000.

[31] A. Ecker and W. Zulehner. On the smoothing property of multi-grid methods in the non-symmetric case. *Numerical Linear Algebra with Applications*, 3(2):161–172, 1996.

[32] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

[33] H. Eschenauer and N. Olhoff. Topology optimization of continuum structures: A review. *Applied Mechanics Rev*, 54(4), 2001.

[34] G. Haase, E. Lindner, W. Mühlhuber, and C. Rathberger. Optimal sizing and shape optimization in structural mechanics. In M. C. J. et al, editor, *Industrial Mathematics*, pages 221–240. Narosa Publishing House, New Delhi, India, 2006.

[35] G. Haase, E. Lindner, and C. Rathberger. Fast shape design for industrial components. In A. D. Bucchianico, R. M. M. Mattheij, and M. A. Peletiers, editors, *Progress in Industrial Mathematics at ECMI 2004*, volume 8 of *European Consortium for Mathematics in Industry*, pages 361–365, Berlin, Heidelberg, 2006. Springer Verlag.

[36] W. Hackbusch. Fast solution of elliptic control problems. *J. Optimization Theory Appl.*, 31:565–581, 1980.

[37] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer-Verlag, Berlin, 1985.

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

[39] J. Haslinger and P. Neittaanmäki. *Finite Element Approximation for Optimal Shape, Material and Topology Design*. John Wiley & Sons Ltd., Chinchester, 2nd edition, 1997.

[40] S. Hazra and V. Schulz. Simultaneous pseudo-timestepping for PDE-model based optimization problems. *BIT*, 44(3):457–472, 2004.

[41] R. H. W. Hoppe, S. I. Petrova, and V. H. Schulz. 3D structural optimization in electromagnetics. In M. G. et al., editor, *Proceedings of Domain Decomposition Methods and Applications (Lyon, Oct. 9–12, 2000)*, 2001.

[42] R. H. W. Hoppe, S. I. Petrova, and V. H. Schulz. Topology optimization of conductive media described by Maxwell's equations. *Lect. Notes Comput. Sci.*, pages 414–422, 2001.

[43] J. B. Jacobsen, N. Olhoff, and E. Rønholt. Generalized shape optimization of three–dimensional structures using materials with optimum microstructures. *Mechanics of Materials*, 28(1–4):207–225, 1998.

[44] A. M. Kalamkarov and A. G. Kolpakov. *Analysis, Design and Optimization of Composite Structures*. John Wiley & Sons, New York; London; Sydney, 1997.

[45] B. Kaltenbacher. On the regularizing properties of a full multigrid method for ill-posed problems. *Inverse Problems*, 17:767–788, 2001.

[46] B. Kawohl, O. Pironneau, L. Tartar, and J.-P. Zolesio, editors. *Optimal Shape Design*, volume 1740 of *Lecture Notes in Mathematics*. Springer, Berlin, New York, Heidelberg, 2000.

[47] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*, volume 120 of *Springer Series in Applied Mathematical Sciences*. Springer, Berlin, 1996.

[48] M. Křížek and P. Neittaanmäki. *Mathematical and Numerical Modelling in Electrical Engineering*. Kluwer Academic Publishers, Dordrecht, 1996.

[49] D. Lukáš. Shape optimization of homogeneous electromagnets. In U. van Rienen, M. Günther, and D. Hecht, editors, *Scientific Computing in Electrical Engineering*, volume 18 of *Lecture Notes in Computational Science and Engineering*, pages 145–152. Springer, 2001.

[50] D. Lukáš. On solution to an optimal shape design problem in 3–dimensional linear magnetostatics. *Applications of Mathematics*, 49(5):441–464, 2004.

[51] D. Lukas. An integration of optimal topology and shape design for magnetostatics. In A. Anile, G. Ali, and G. Mascali, editors, *Proceedings of SCEE 2004*, Mathematics in Industry 9, pages 227–232. Springer-Verlag, 2006.

[52] D. Lukas. Multigrid-based optimal shape and topology design in magnetostatics. In T. Boyanov, S. Dimova, K. Georgiev, and G. Nikolov, editors, *Proceedings of NM&A 2006*, pages 82–91, Berlin, Heidelberg, 2006. Springer-Verlag.

[53] D. Lukas. A multigrid method for coupled optimal topology and shape design in nonlinear magnetostatic. In A. B. de Castro, D. Gomes, P. Quintela, and P. Salgado, editors, *Proceedings of ENUMATH 2005*, pages 1015–1022, Berlin, Heidelberg, 2006. Springer-Verlag.

[54] D. Lukáš. *Optimal Shape Design in Magnetostatics*. PhD thesis, VŠB–Technical University Ostrava, Czech Republic, September 2003.

[55] D. Lukáš and P. Chalmoviansky. A sequentila coupling of optimal topology and multilevel shape design applied to 2-dimensional nonlinear magnetostatics. *Computing and Visualization in Science*, 10(3):135–144, 2007.

[56] D. Lukas, U. Langer, E. Lindner, R. Stainko, and J. Pistora. Computational shape and topology optimization with applications to 3-dimensional magnetostatics. In *Proceedings of the Oberwolfach Workshop on Computational Electromagnetism 2004*, volume 1, pages 601–603, 2004.

[57] B. Maar and V. Schulz. Interior point multigrid methods for topology optimization. *Structural and Multidisciplinary Optimization*, 19:214–224, 2000.

[58] A. Marrocco and O. Pironneau. Optimum design with Lagrangian finite elements: Design of an electromagnet. *Comp. Meth. Appl. Mech. Eng.*, 15:277–308, 1978.

[59] L. Modica and S. Mortola. Un esempio di Γ-convergenza. *Boll. Unione Mat. Ital.*, 14(B):285–299, 1977.

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

[61] F. Murat and J. Simon. *Studies on Optimal Shape Design Problems*, volume 41 of *Lecture Notes in Computational Science*. Springer, Berlin, 1976.

[62] P. Neittaanmäki, M. Rudnicki, and A. Savini. *Inverse Problems and Optimal Design in Electricity and Magnetism*. Oxford University Press, 1995.

[63] V. Nübel, A. Düster, and E. Rank. An *rp*–adaptive finite element method for elastoplasic problems. In *ECCOMAS 2004*, 2004.

[64] N. Olhoff and J. E. Taylor. On structural optimization. *Journal of Applied Mathematics*, 50(4):1139–1151, 1983.

[65] C. Pechstein and B. Jüttler. Monotonicity-preserving interproximation of B-H-curves. *Journal of Computational and Applied Mathematics*, 196(1):45–57, 2006.

[66] P. Pedersen, editor. *Optimal Design with Advanced Materials*. Elsevier, Amsterdam, The Netherlands, 1993.

[67] G. H. Peichl and W. Ring. Optimization of the shape of an electromagnet: Regularity results. *Advances in Mathematical Sciences and Applications*, 8:997–1014, 1998.

[68] G. H. Peichl and W. Ring. Asymptotic commutativity of differentiation and discretization in shape optimization. *Mathematical and Computing Modelling*, 29:19–37, 1999.

[69] J. Petersson. On continuity of the design–to–state mappings for trusses with variable topology. *International Journal of Engineering Science*, 39(10):1119–1141, 2001.

[70] J. Petersson and J. Haslinger. An approximation theory for optimum sheets in unilateral contact. *Quarterly of Applied Mathematics*, 56(2):309–332, 1998.

[71] J. Petersson and O. Sigmund. Slope constrained topology optimization. *Internat. J. Numer. Methods Engrg.*, 41:1417–1434, 1998.

[72] O. Pironneau. *Optimal Shape Design for Elliptic Systems*. Springer series in Computational Physics. Springer, New-York, 1984.

[73] E. Ramm, K. Maute, and S. Schwarz. Adaptive topology and shape optimization. *Computational Mechanics*, 1998.

[74] J. Rasmussen, M. Damsgaard, S. T. Christensen, and E. Surma. Design optimization with respect to ergonomic properties. *Structural and Multidisciplinary Optimization*, 24:89–97, 2002.

[75] A. Rietz. Sufficiency of a finite exponent in the simp (power law) method. *Structural and Multidisciplinary Optimization*, 12:159–163, 2001.

[76] A. Rösch and R. Simon. Linear and discontinuous approximations for optimal control problems. *Numerical Functional Analysis and Optimization*, 26(3):427–448, 2005.

[77] A. Rösch and R. Simon. Superconvergence properties for optimal control problems discretized by piecewise linear and discontinuous functions. *Numerical Functional Analysis and Optimization*, 28(3-4):425–443, 2007.

[78] G. I. N. Rozvany, editor. *Shape and Lay–Out Optimization of Structural Systems and Optimality Criteria Methods*. CISM Lecture Notes no. 325. Springer, Wien, 1992.

[79] G. I. N. Rozvany, editor. *Optimization of Large Structural Systems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[80] G. I. N. Rozvany. Aims, scope, methods, history and unified terminology of computer–aided topology optimization in structural mechanics. *Structural and Multidisciplinary Optimization*, 21:90–108, 2001.

[81] M. Save and W. Prager, editors. *Structural Optimization, Vol. 1, Optimality Criteria*. Plenum Press, New York, 1985.

[82] M. Save and W. Prager, editors. *Structural Optimization, Vol. 2, Mathematical Optimization*. Plenum Press, New York, 1990.

[83] O. Scherzer. An iterative multi level algorithm for solving nonlinear ill–posed problems. *Numer. Math.*, 80:579–600, 1998.

[84] A. Schleupen, K. Maute, and E. Ramm. Adaptive FE–procedures in shape optimization. *Structural and Multidisciplinary Optimization*, 19:282–302, 2000.

[85] J. Schöberl and W. Zulehner. On Schwarz-type smoothers for saddle point problems. *Numerische Mathematik*, 95(2):377–399, 2003.

[86] J. Schöberl and W. Zulehner. Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems. SFB-Report 2006-19, Johannes Kepler University Linz, SFB "Numerical and Symbolic Scientific Computing", 2006.

[87] J. Schöberl and W. Zulehner. Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems. *SIAM J. Matrix Anal. Appl.*, 29:752–773, 2007.

[88] V. H. Schulz. Incomplete indefinite decompositions as multigrid smoothers for KKT systems. *International Series of Numerical Mathematics*, 133:257–266, 1999.

[89] M. Shallaby. *Spline Implicitization of Planar Shapes and Applications*. PhD thesis, RISC, University Linz, July 2003.

[90] O. Sigmund. *Design of material structures using topology optimization*. PhD thesis, Technical University of Denmark, Department of Solid Mechanics, 1994.

[91] O. Sigmund and J. Petersson. Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Structural and Multidisciplinary Optimization*, 16:68–75, 1998.

[92] R. Simon and W. Zulehner. On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems. Technical Report 01, Special Research Program SFB F013, JK University of Linz, January 2007.

[93] J. Sokolowski and J.-P. Zolesio. *Introduction to Shape Optimization*. Number 16 in Springer Series in Computational Mathematics. Springer, Berlin, 1992.

[94] W. Stadler. Non–existence of solutions in optimal structural design. *Opt. Control Appl. Meth.*, 7:243–258, 1986.

[95] R. Stainko. An adaptive multilevel approach to stress constrained topology optimization problems. Technical Report 2004-20, J. Kepler University Linz, SFB F013, 2004.

[96] R. Stainko. An adaptive multilevel approach to minimal compliance topology optimization. *Communications in Numerical Methods in Engineering*, 22:109–118, 2006.

[97] R. Stainko. An adaptive multilevel approach to the minimal compliance problem in topology optimization. *Communications in Numerical Methods in Engineering*, 22:109–118, 2006.

[98] R. Stainko. *Advanced Multilevel Techniques to Topology Optimization*. PhD thesis, Johannes Kepler University Linz, Austria, Special Reasearch Project SFB F013, 2006.

[99] R. Stainko. An optimal solver to a kkt-system. Technical Report 13, Special Research Program SFB F013, JK University of Linz, March 2006.

[100] R. Stainko and M. Burger. A one shot approach to topology optimization with local stress constraints. In M. P. Bendsoe, N. Olhoff, and O. Sigmund, editors, *IUATM Symposium on Topology Design Optimization of Structures, Machines and Materials*, volume 137 of *Solid Mechanics and its Applications*, pages 181–184. IUTAM, Springer, 2006.

[101] M. Stolpe and K. Svanberg. An alternative interpolation scheme for minimum compliance topology optimization. *Structural and Multidisciplinary Optimization*, 22:116–124, 2001.

[102] M. Stolpe and K. Svanberg. Modeling topology optimization problems as linear mixed 0–1 programs. *International Journal for Numerical Methods in Engineering*, 57(5):723–739, 2003.

[103] K. Svanberg. The method of moving asymptotes - a new method for structural optimization. *International Journal for Numerical Methods in Engineering*, 24:359–373, 1987.

[104] K. Svanberg. A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM Journal of Optimization*, 12(2):555–573, 2002.

[105] S. Ta'asan. "One-shot" methods for optimal control of distributed parameter systems I: The finite dimensional case. ICASE-Report 91-2, NASA Langley Research Center, Hampton VA, 1991.

[106] N. Takahashi. Optimization of die press model. In *Proceedings of the TEAM Workshop in the Sixth Round*, Okayama, Japan, March 1996.

[107] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen. Theorie, Verfahren und Anwendungen*. Vieweg, Wiesbaden, 2005.

[108] S. Vanka. Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *J. Comput. Phys.*, 65:138–158, 1986.

[109] A. Waechter and L. Biegler. On the implementation of an interior–point filter line–search algorithm for large–scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

[110] Y. M. Xie and G. P. Steven. *Evolutionary Structural Optimization*. Springer, Berlin, 1997.

[111] J. Yoo and N. Kikuchi. Topology optimization in magnetic fields using the homogenization design method. *Int. J. Numer. Meth. Engng.*, 48:1463–1479, 2000.

[112] K. Yuge and N. Kikuchi. Optimization of a frame structure subjected to a plastic deformation. *Structural Optimization*, 10(3–4):197–208, 1995.

[113] W. Zulehner. A class of smoothers for saddle point problems. *Computing*, 65:227–246, 2000.

[114] W. Zulehner. Analysis of iterative methods for saddle point problems: A unified approach. *Mathematics of Computation*, 71(238):479–505, 2001.