On a Multigrid Preconditioned Augmented Lagrangians Applied to the Stokes and Optimization Problems

PANM 2006 in honor of Prof. Babuška, May 30, Prague

D. Lukáš and Z. Dostál

Department of Applied Mathematics,
VŠB–Technical University Ostrava

email: dalibor.lukas@vsb.cz

Outline

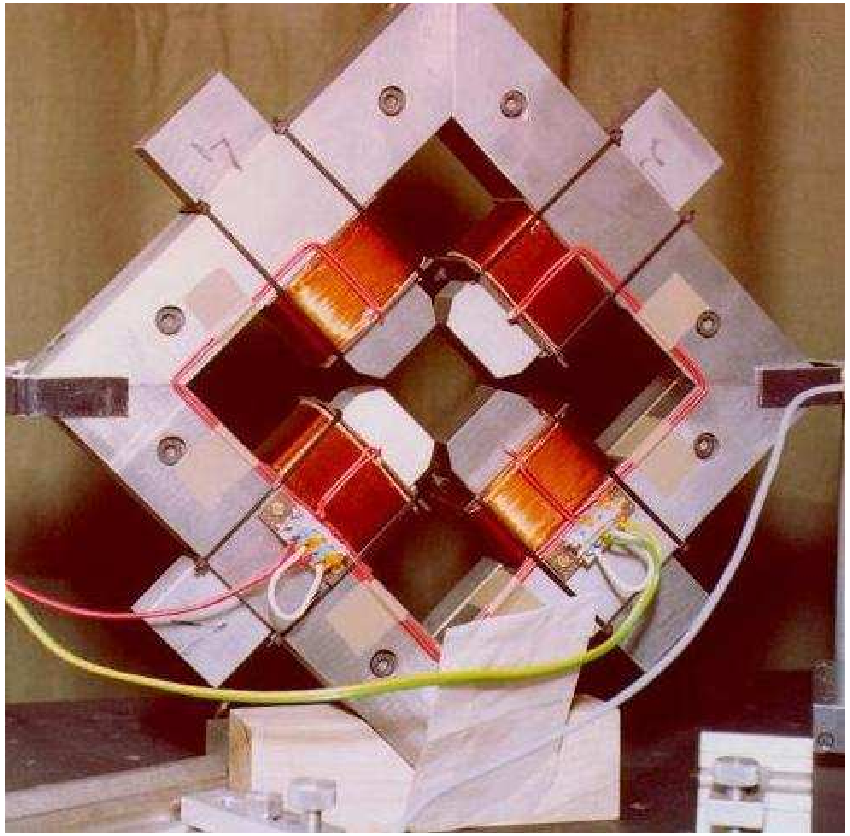
- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian method
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian method
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

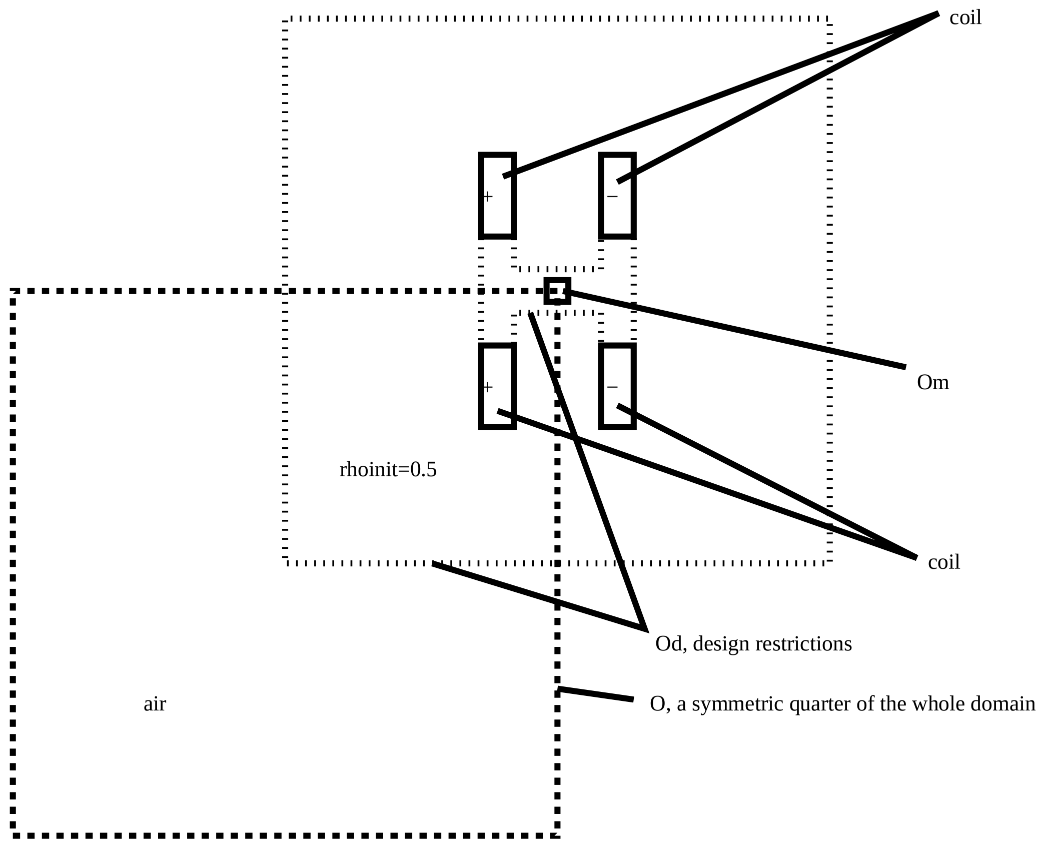
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



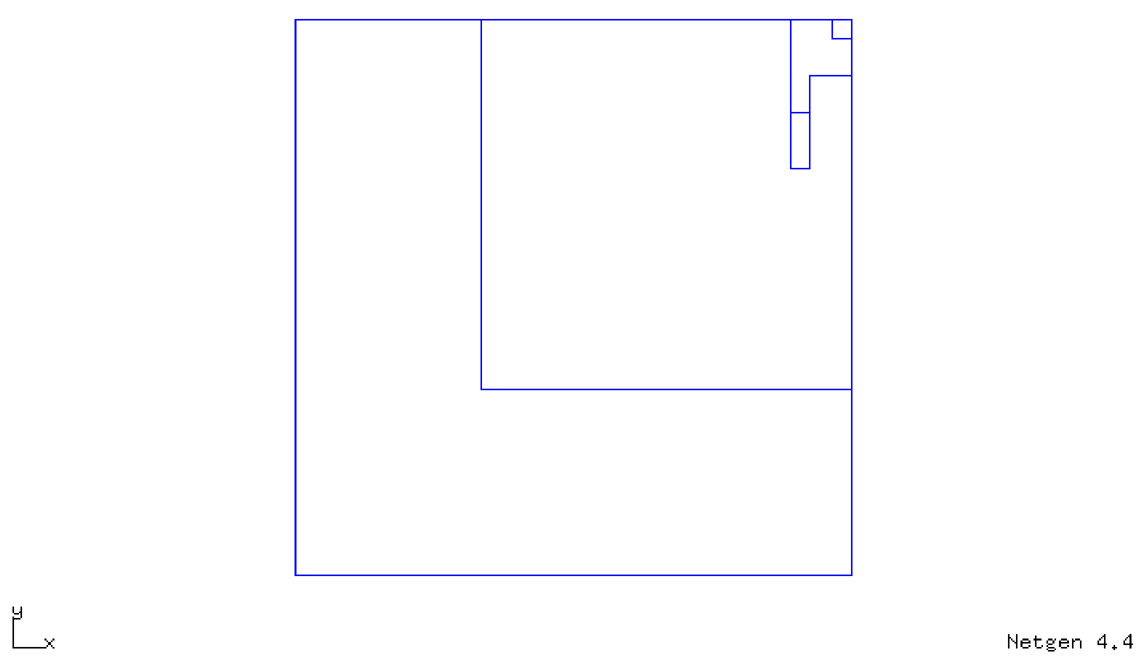
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



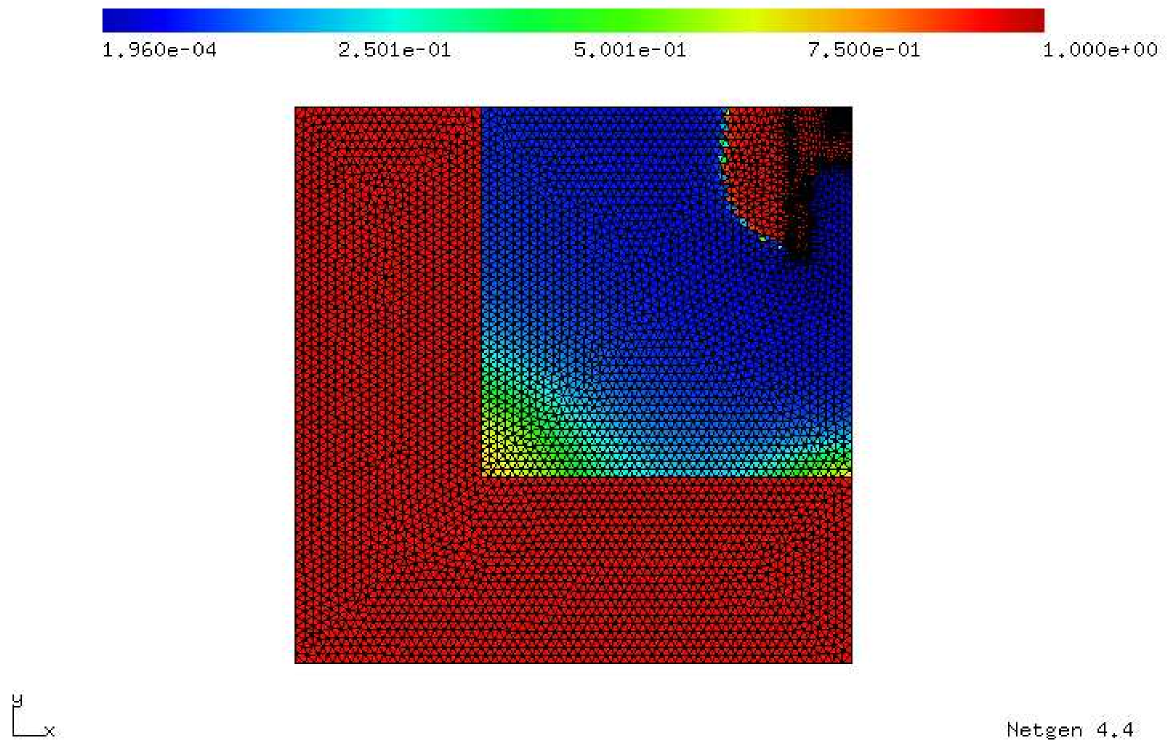
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



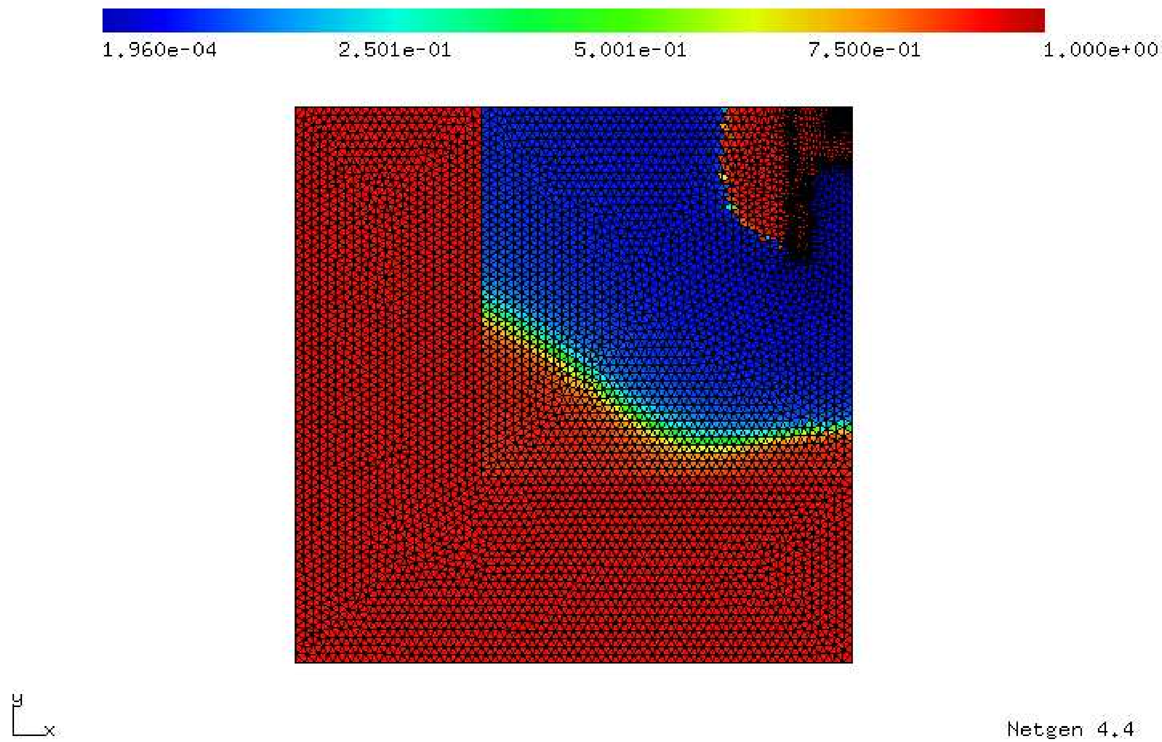
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



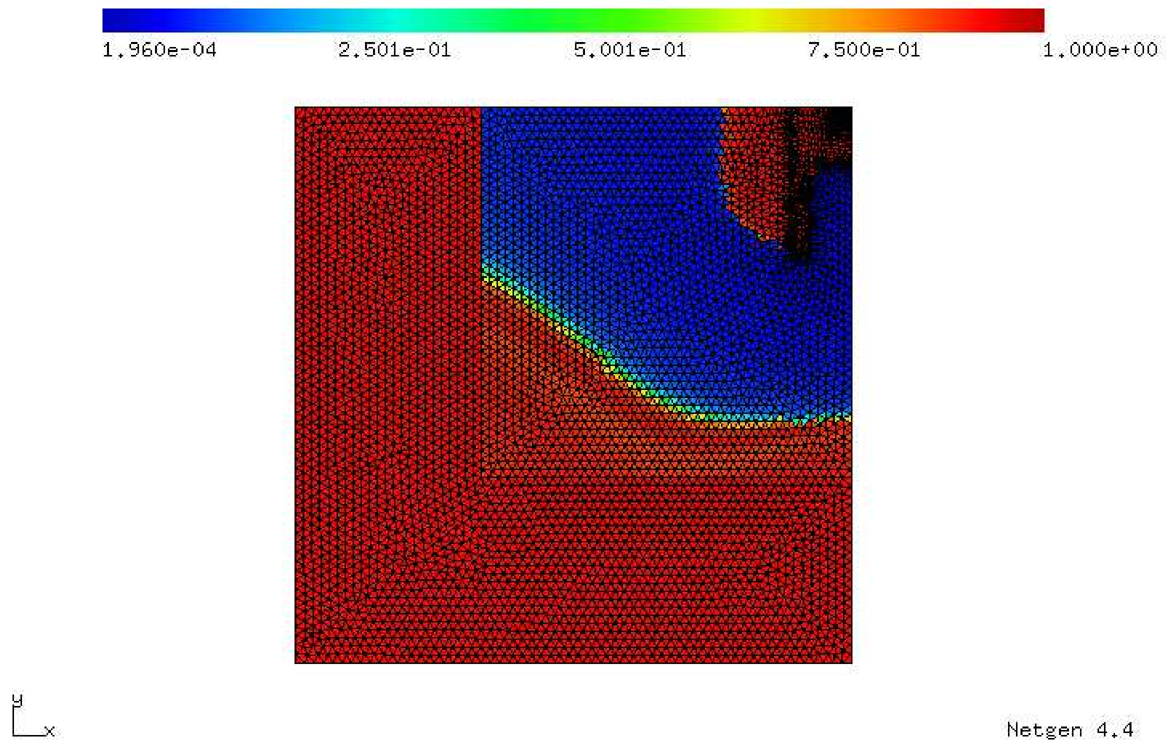
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



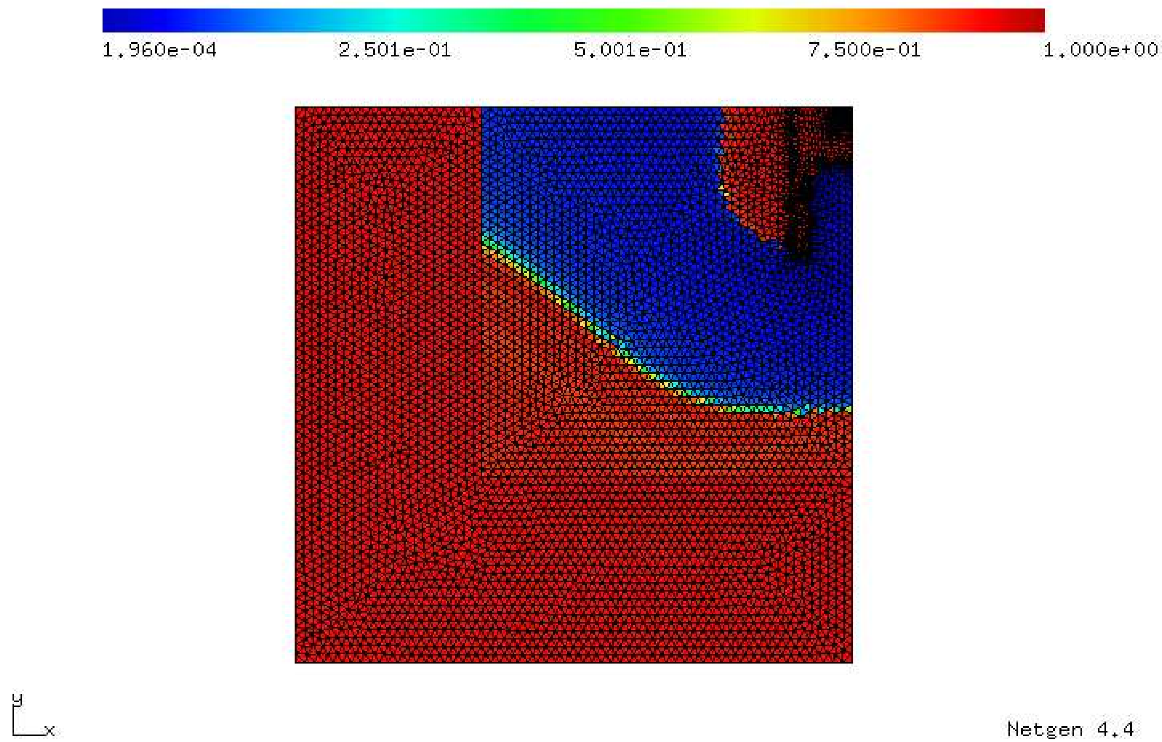
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



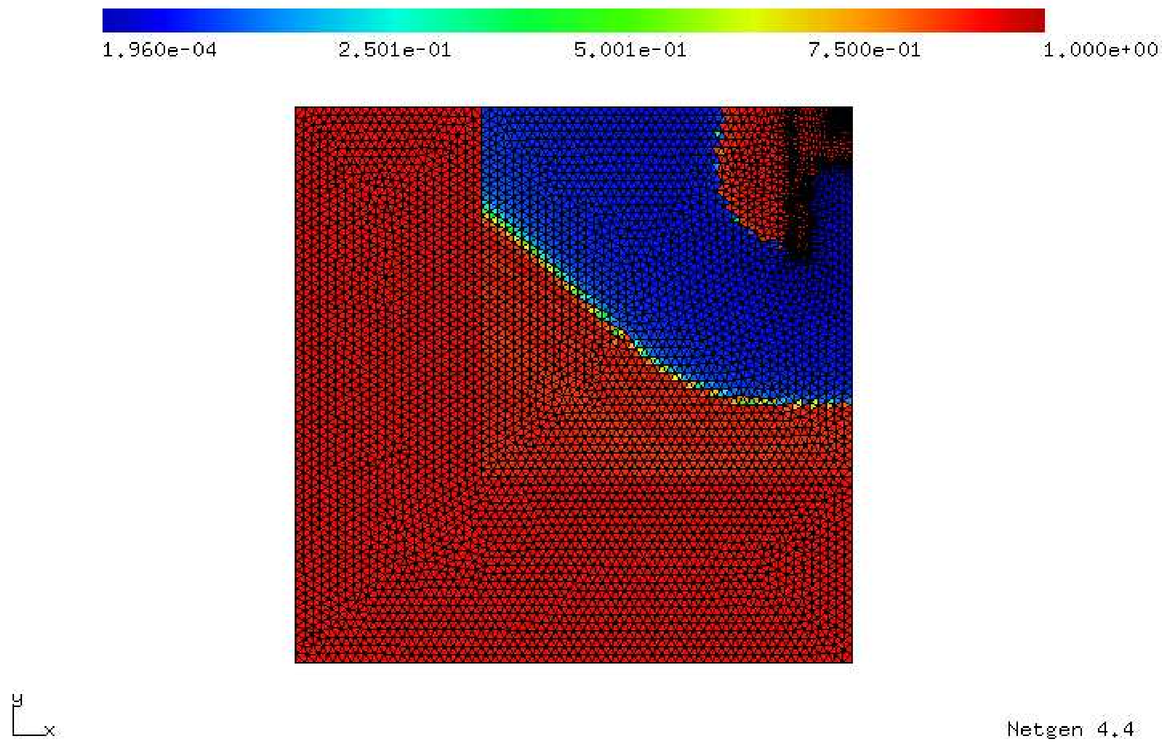
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



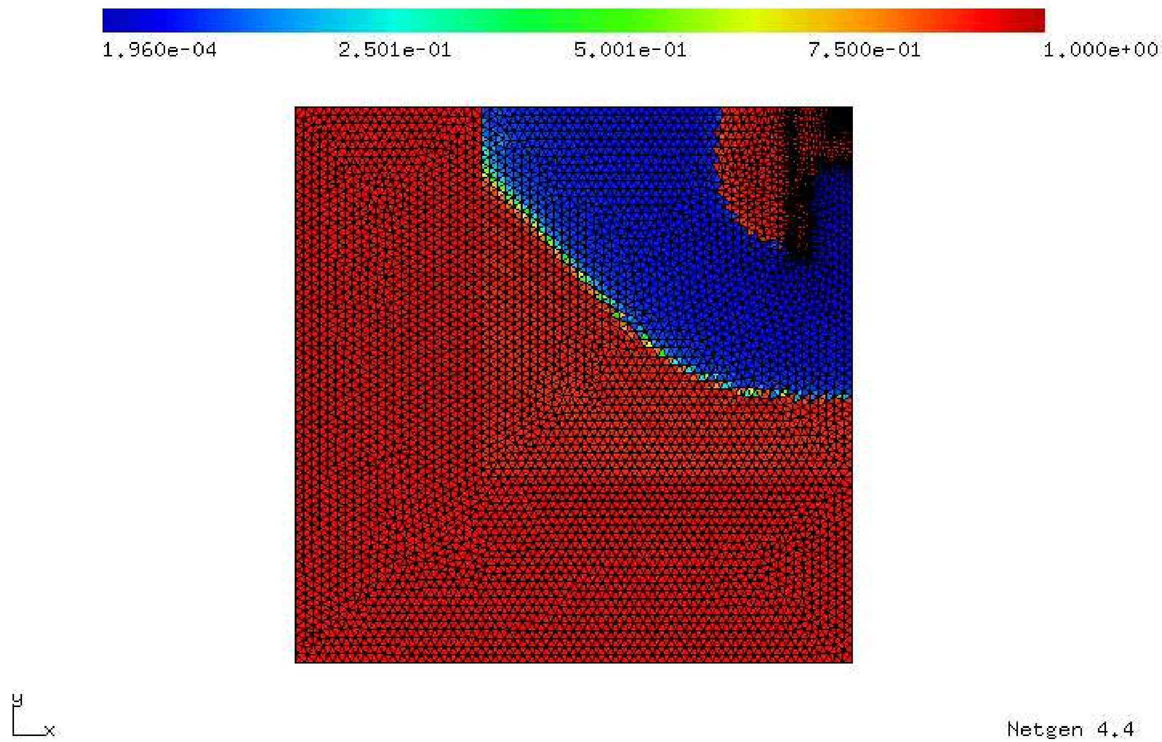
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



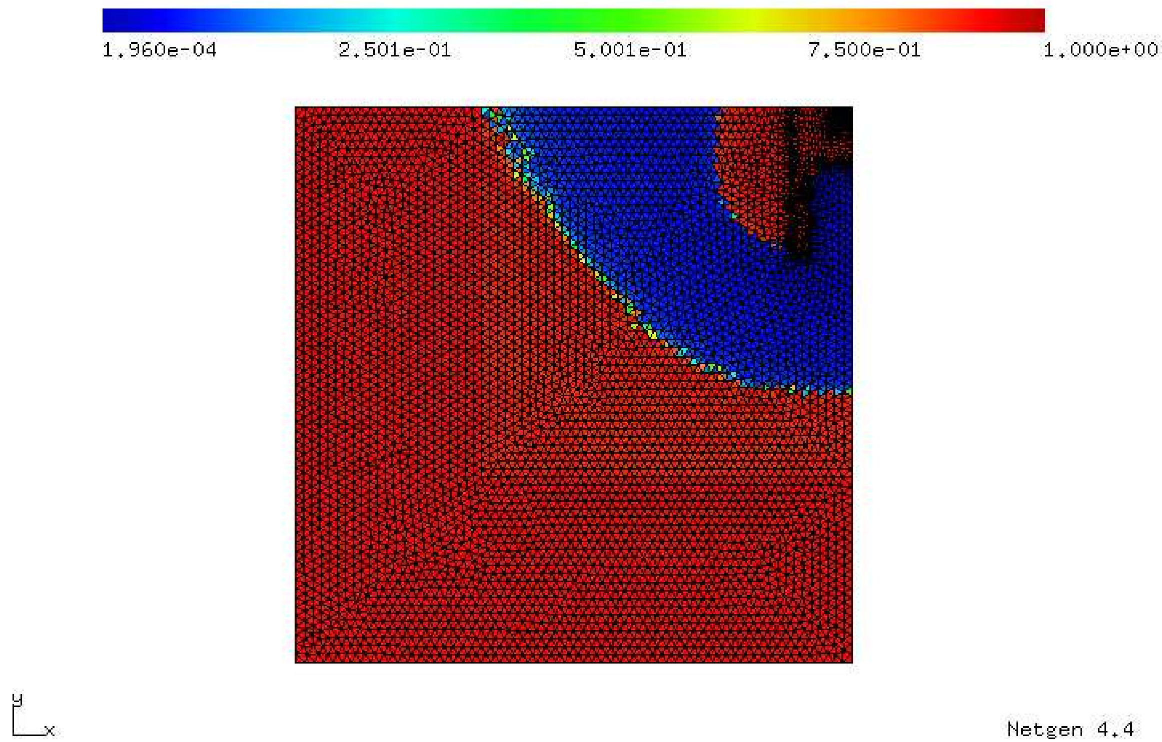
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



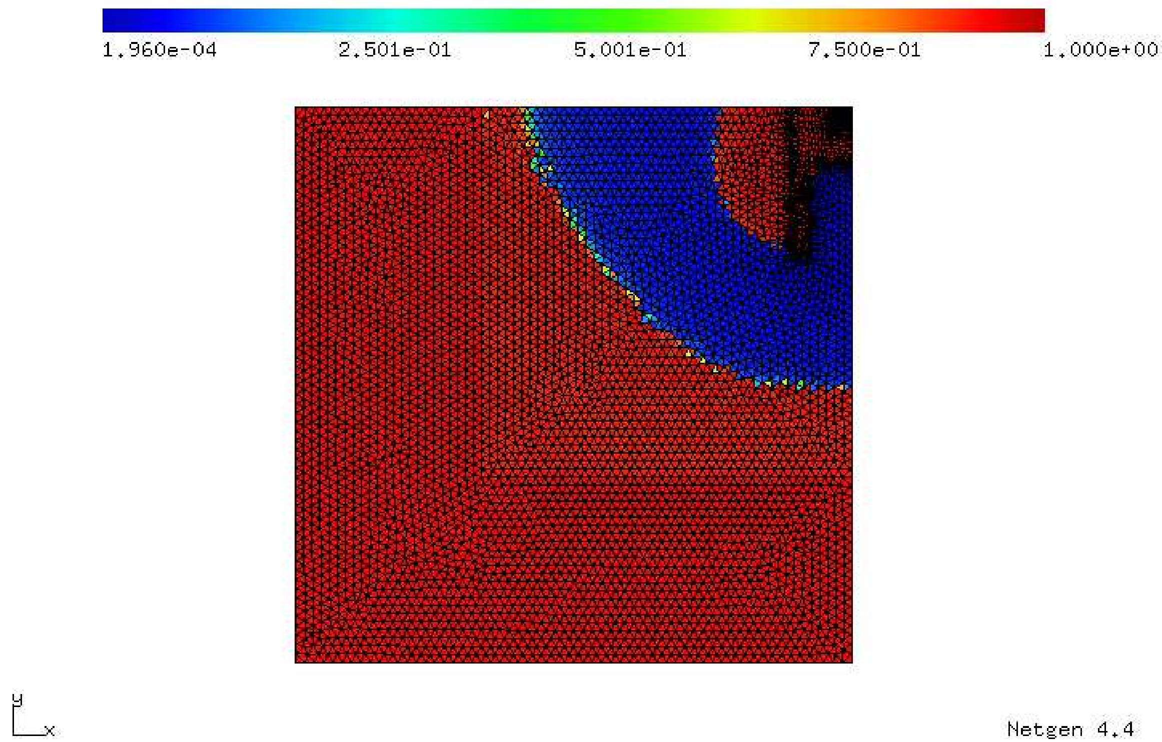
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



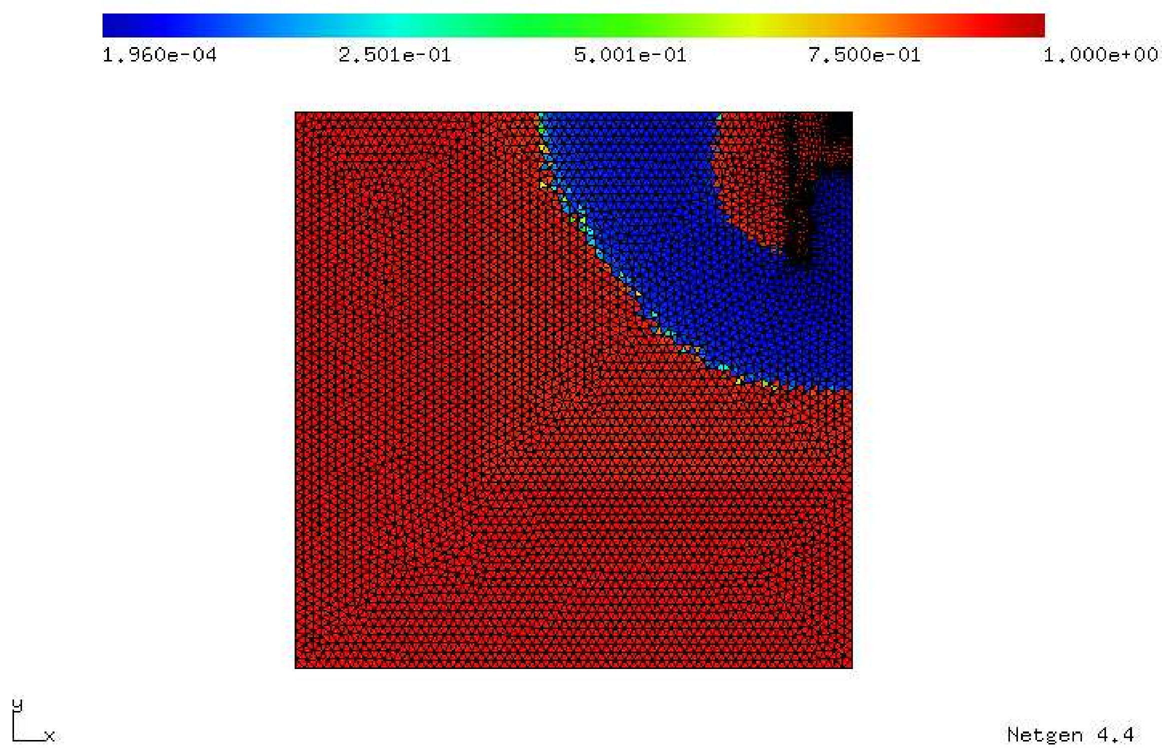
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



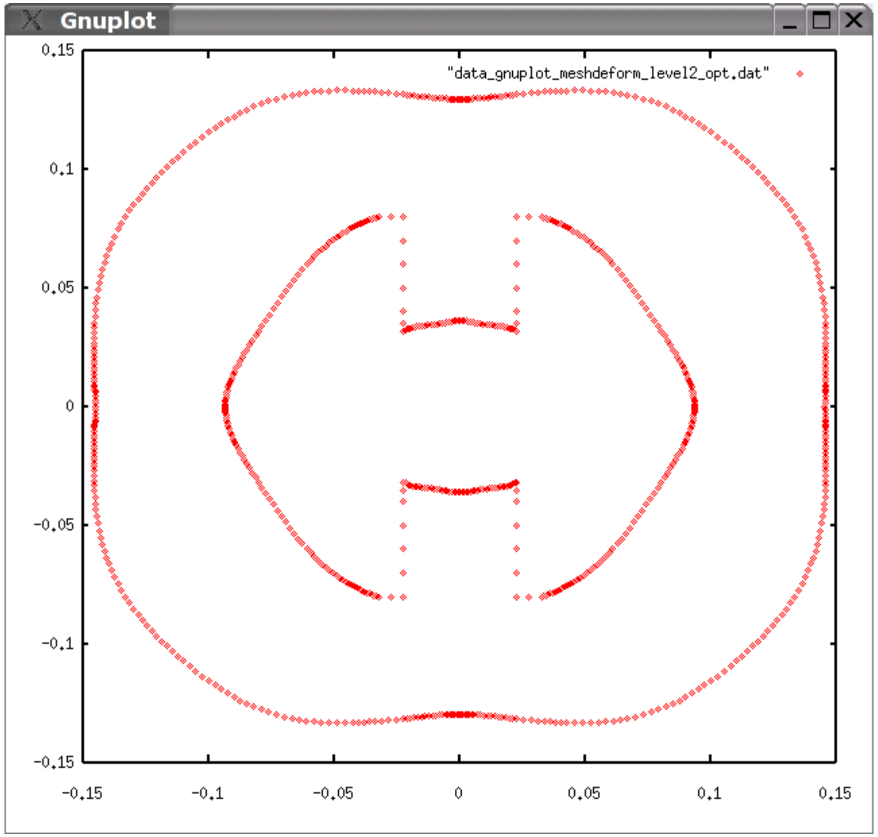
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



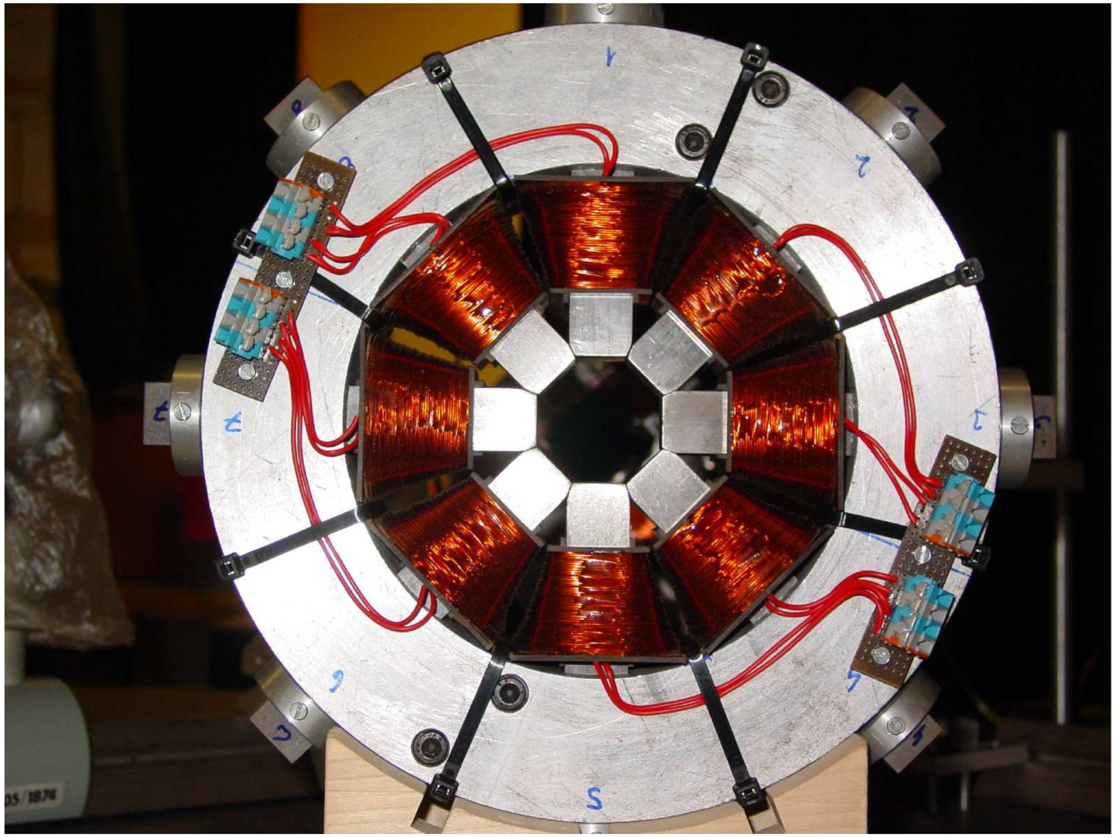
Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



Motivation: Large–scale discretized topology optimization

Topology optimization of an electromagnet



Motivation: Large–scale discretized topology optimization

Topology optimization

$$\min_{q\in Q} J(q,u) \quad \text{subject to} \quad B(q)u = g \text{ na } V'$$

In order to use sequential quadratic programming, we need the Hessian linearization of the functional J.

State elimination

$$\min_{q\in Q}\; J(q, B(q)^{-1}g)$$

Discretizations of the linearized Hessian is large–scale and dense!

Simultaneous (all–at–once) method

$$\min_{(q,u)\in Q\times V}\;\max_{\lambda\in V}\{J(q,u) + \langle B(q)u - g, \lambda\rangle_V\}$$

Discretizations of the linearized Hessian are even bigger, but sparse and well–structured!

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian method
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Setting of the Saddle–Point Problem

Saddle–point problem

V, Z …Hilbert spaces,
$(.,.)_V$, $(.,.)_Z$ …inner products, $\langle .,. \rangle_V$, $\langle .,. \rangle_Z$ …duality pairings:

$$\min_{u \in V} h(u) \quad \text{s.t. } Bu = g,$$

where $h(u) := \frac{1}{2}\langle Au, u\rangle_V - \langle f, u\rangle_V$, $A : V \to V'$ linear bounded self–adjoint and positive definite, $B : V \to Z'$ linear and bounded, $f \in V'$, $g \in \text{Range}(B)$.

Lagrange formalism

$p \in Z$ …Lagrange multiplier

$$\min_{u \in V} \max_{p \in Z} \{h(u) + \langle Bu - g, p\rangle_Z\} \quad \text{equivalent to} \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

where $B^T : Z \to V'$ adjoint to B, i.e. $\langle B^T z, v\rangle_V = \langle Bv, z\rangle_Z$.
$\exists! u^* \in V$, $\exists p^* \in Z$. Moreover, if $\text{Kernel}(B^T) = \{0\}$, then $\exists! p^* \in Z$.

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian method
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Optimal Uzawa–Type Methods

Symmetric Uzawa methods

$\widehat{A}^{-1}$ ... preconditioner to A, $\widehat{S}^{-1}$ ... preconditioner to Schur complement $S := BA^{-1}B^T$,
α_V, α_Z ... small enough positive damping parameters.

The outer k–th iteration:

$$\begin{aligned} u^{(k+1/2)} &:= u^{(k)} + \alpha_V \widehat{A}^{-1}\left[f - \left(Au^{(k)} + B^T p^{(k)}\right)\right], \\ p^{(k+1)} &:= p^{(k)} + \alpha_Z \widehat{S}^{-1}\left[Bu^{(k+1/2)} - g\right], \\ u^{(k+1)} &:= u^{(k)} + \alpha_V \widehat{A}^{-1}\left[f - \left(Au^{(k)} + B^T p^{(k+1)}\right)\right]. \end{aligned}$$

α_V a α_Z often difficult to find!

The Uzawa method is the outher loop and the inner loop is preconditioning using multigrid and/or PCG.

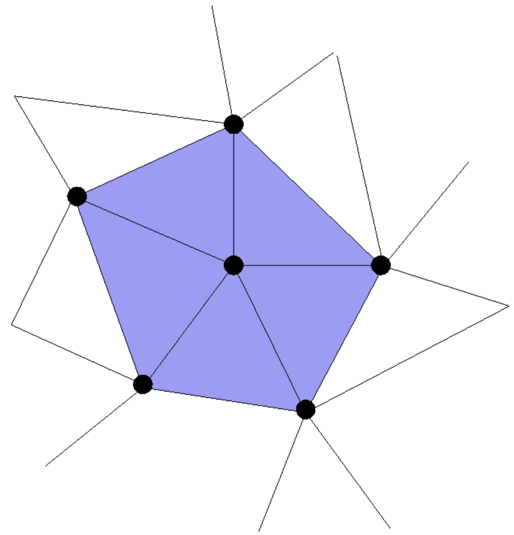
Optimal Uzawa–Type Methods

Schöberl & Zulehner, 2003: Additive multigrid smoother

(V_i), (Z_i) …sequences of FE-subspaces of V and Z, respectively; (P_i), (Q_i) …the corresponding prolongations, i.e. $P_i : V_i \to V$, $Q_i : Z_i \to Z$:

$$u^{(k+1)} := u^{(k)} + \alpha_V \sum_{i=1}^{N} P_i w_i^{(k)},$$

$$p^{(k+1)} := p^{(k)} + \alpha_Z \sum_{i=1}^{N} Q_i r_i^{(k)},$$



where $(w_i^{(k)}, r_i^{(k)})$ is a solution to a local saddle–point subproblem, typically defined around a discret. node:

$$\begin{pmatrix} P_i^T A P_i & P_i^T B^T Q_i \\ Q_i^T B P_i & 0 \end{pmatrix} \begin{pmatrix} w^{(k+1)} \\ r^{(k+1)} \end{pmatrix} = \begin{pmatrix} P_i^T \left[f - A u^{(k)} \right] \\ Q_i^T \left[g - B u^{(k)} \right] \end{pmatrix}.$$

Equivalent to a symmetric Uzawa method.

Multigrid methods

Vanka, 1986: Multiplicative smoother

(V_i), (Z_i) …sequences of FE-subspaces of V and Z; (P_i), (Q_i) …the corresponding prolongations, i.e. $P_i : V_i \to V$, $Q_i : Z_i \to Z$:

$$u^{(k+\frac{i}{N})} := u^{(k+\frac{i-1}{N})} + P_i w_i^{(k+\frac{i}{N})},$$
$$p^{(k+\frac{i}{N})} := p^{(k+\frac{i-1}{N})} + Q_i r_i^{(k+\frac{i}{N})},$$

where

$$\begin{pmatrix} P_i^T A P_i & P_i^T B^T Q_i \\ Q_i^T B P_i & 0 \end{pmatrix} \begin{pmatrix} w^{(k+\frac{i}{N})} \\ r^{(k+\frac{i}{N})} \end{pmatrix} = \begin{pmatrix} P_i^T \left[f - A u^{(k+\frac{i-1}{N})} - B^T p^{(k+\frac{i-1}{N})} \right] \\ Q_i^T \left[g - B u^{(k+\frac{i-1}{N})} \right] \end{pmatrix}.$$

The multiplicative smoother is numerically more efficient than additive, no damping parameters needed, but for saddle–point problems there is no convergence theory!

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian methods
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Augmented Lagrangians Methods

$h(u) := (1/2)\langle Au, u\rangle_V - \langle f, u\rangle_V$, $\rho \ldots$ positive augmented Lagrange penalty, $\iota_Z : Z \to Z'$ inner product on Z (Riesz isomorphism).

The outer k-th iteration:

Step 1. $u^{(k+1)}$ solves $\min\limits_{u \in V} \left\{ h(u) + \langle Bu - g, p^{(k)}\rangle_Z + \frac{\rho^{(k)}}{2} \|Bu - g\|_{Z'}^2 \right\}$,

Step 2. $p^{(k+1)} := p^{(k)} + \rho^{(k)} \iota_Z^{-1} \left(Bu^{(k+1)} - g \right)$,

Step 3. Increase $\rho^{(k)}$.

Denote:
$L(u, p, \rho) := h(u) + \langle Bu - g, p\rangle_Z + \frac{\rho}{2}\|Bu - g\|_{Z'}^2$ Lagrange functional, $F(u, p, \rho) := \nabla_u L(u, p, \rho) = Au - f + B^T p + \rho B^T \iota_Z^{-1}(Bu - g)$ its Fréchet derivative.

Then, Step 1 is equivalent to

$$\left[A + \rho^{(k)} B^T \iota_Z^{-1} B \right] u^{(k+1)} = f - B^T (p^{(k)} - \iota_Z^{-1} g).$$

Semi–Monotonic Augmented Lagrangians Method

Dostál, 2005: The algorithm

The k-th iteration:

Find $u^{(k+1)} : \|F(u^{(k+1)}, p^{(k)}, \rho^{(k)})\|_{V'} \leq \min\left\{\nu \|Bu^{(k+1)} - g\|_{Z'}, \eta\right\}$ %PCG◯◯MG

if $\|F(u^{(k+1)}, p^{(k)}, \rho^{(k)})\|_{V'} \leq \varepsilon$ and $\|Bu^{(k+1)} - g\|_{Z'} \leq \varepsilon_{\text{feas}}$ **then**

  break

end if

$p^{(k+1)} := p^{(k)} + \rho^{(k)} \iota_Z^{-1}(Bu^{(k+1)} - g)$

if $L(u^{(k+1)}, p^{(k+1)}, \rho^{(k)}) < L(u^{(k)}, p^{(k)}, \rho^{(k)}) + \frac{\rho^{(k)}}{2}\|Bu^{(k+1)} - g\|_{Z'}$ **then**

  $\rho^{(k+1)} := \beta \rho^{(k)}$

else

  $\rho^{(k+1)} := \rho^{(k)}$

end if

$\nu > 0$, $\eta > 0$, $\beta > 1$, $\varepsilon > 0$, $\varepsilon_{\text{feas}} > 0$ and where $\|Bu - g\|_{Z'} := \sqrt{\left\langle Bu - g, \iota_Z^{-1}(Bu - g)\right\rangle_Z}$ a $\|F(u,p,\rho)\|_{V'} := \sqrt{\left\langle F(u,p,\rho), \iota_V^{-1} F(u,p,\rho)\right\rangle_V}$.

Semi–Monotonic Augmented Lagrangians Method

Dostál, 2005: Analysis

1. The number of the outer iterations is bounded:

$$k \leq \left(\frac{(1+s)\eta^2 + \|f\|_{V'}^2}{\lambda \rho^{(0)}} + \rho^{(0)} \|g\|_{Z'}^2 \right) / \min{}^2 \{\nu\varepsilon, \varepsilon_{\text{feas}}\} =: k_{\max},$$

where $\lambda > 0$ is the smallest eigenvalue of A, $s > 0$ is the smallest integer so that $\beta^s \rho^{(0)} \geq \frac{\nu^2}{\lambda}$.

2. The augmented Lagrange parameter is bounded

$$\rho^{(k)} \leq \beta^s \rho^{(0)} < \frac{\beta \nu^2}{\lambda} =: \rho_{\max}.$$

Consequently, we have the uniform spectral equivalence

$$\forall k : A + \rho^{(k)} B^T M^{-1} B \approx A,$$

since $\forall v \in V : \lambda(v,v)_V \leq \left(\|A\|_{V'}^2 + \rho_{\max} \|B^T M^{-1} B\|_{V'}^2 \| \right) (v,v)_V$ and $\|v\|_A \approx \|v\|_V$.

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian methods
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Multigrid Preconditioned Augmented Lagrangians Method

Implementation

requires only the following changes:

- In PCG (step 1) we use the preconditioner $\widehat{A}^{-1}$.
- Each occurences of ι_V^{-1} and ι_Z^{-1} are replaced by $\widehat{A}^{-1}$ a $\widehat{M}^{-1}$, respectively.
- The norms are calculated as follows:

$$\|Bu-g\|_{\widehat{Z}'} := \sqrt{(Bu-g)^T \cdot \widehat{M}^{-1} \cdot (Bu-g)}$$

and

$$\|\widetilde{F}(u,p,\widehat{\rho})\|_{\widehat{V}'} := \sqrt{\widetilde{F}(u,p,\widehat{\rho})^T \cdot \widehat{A}^{-1} \cdot \widetilde{F}(u,p,\widehat{\rho})}.$$

Then, $\lambda_{\min}\left(\widehat{A}_{\frac{1}{2}}^{-1} A \widehat{A}_{\frac{1}{2}}^{-T}\right) \approx 1$, $\operatorname{cond}\left(\widehat{A}_{\frac{1}{2}}^{-1} A \widehat{A}_{\frac{1}{2}}^{-T}\right) \approx 1$ yields an optimal method.

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian methods
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

An Application to the Stokes Problem

The Stokes problem

$\Omega \subset \mathbb{R}^2$ ... bounded polygonal computational domain, $\mathbf{f} \in [L^2(\Omega)]^2$

$$\begin{aligned} -\triangle \mathbf{u} + \boldsymbol{\nabla} p = \mathbf{f} &\quad \text{in } \Omega, \\ \operatorname{div} \mathbf{u} = 0 &\quad \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} &\quad \text{on } \partial\Omega. \end{aligned}$$

Fitting to our framework

$$V := [H_0^1(\Omega)]^2, \quad Z := L^2(\Omega), \quad \langle A\mathbf{u}, \mathbf{v}\rangle_V := \int_\Omega \sum_{i=1}^{2} \boldsymbol{\nabla} u_i \cdot \boldsymbol{\nabla} v_i \, d\mathbf{x},$$

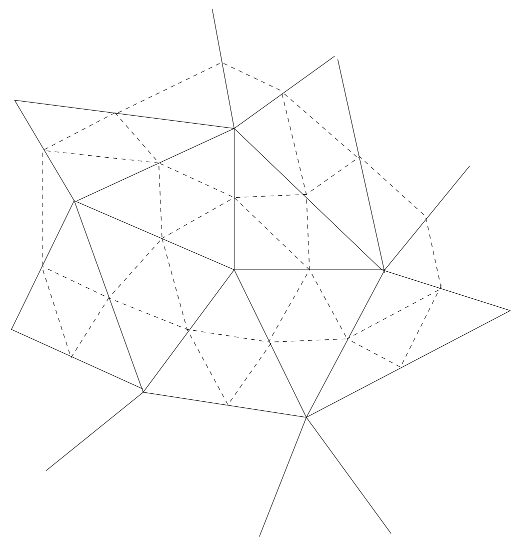
$$\langle b, \mathbf{v}\rangle_V := \int_\Omega \mathbf{f} \cdot \mathbf{v} d\mathbf{x}, \quad \langle B\mathbf{u}, z\rangle_Z := \int_\Omega \operatorname{div} \mathbf{u} \, z \, d\mathbf{x}.$$

Thus, A is a tensor product Laplacian and M is the L^2–inner product.

An Application to the Stokes Problem

FE–multigrid discretization

$(\mathcal{T}_l)$ . . . a nested sequence of triangulations of Ω



Crouzeix–Raviart elements

A nonnested sequence of FE–spaces:

$$V_l := \Big\{ v \in \left[L^2(\Omega)\right]^2 : v|_T \text{ is linear for all } T \in \mathcal{T}_l,$$
$$v \text{ is continuous at the midpoints of interelement boundaries}$$
$$\text{and } v = 0 \text{ along } \partial\Omega \Big\},$$
$$Z_l := \left\{ z \in L^2(\Omega) : z|_T \text{ is constant for all } T \in \mathcal{T}_l \right\}.$$

An Application to the Stokes Problem

Brenner, 1993: Multigrid preconditioning

Smoothers: multilevel diagonal scaling for $\widehat{A_l^{-1}}$, $M_l^{-1} = [\mathrm{diag}(M_l)]^{-1}$

Prolongations: $I_{l-1}^l : V_{l-1} \times Z_{l-1} \to V_l \times Z_l$ given by

$$I_{l-1}^l(v, z) := (J_{l-1}^l v, z)$$

with

$$J_{l-1}^l v(m_e) := \begin{cases} v(m_e) & \text{if } m_e \in \operatorname{int} T \text{ for some } T \in \mathcal{T}_{l-1} \\ \frac{1}{2}\left[v|_{T_1} + v|_{T_2}\right] & \text{if } e \subset T_1 \cap T_2 \text{ for some } T_1, T_2 \in \mathcal{T}_{l-1} \end{cases}$$

at midpoints m_e of internal edges e in $\mathcal{T}_l$.

An Application to the Stokes Problem

Numerical results

$\Omega := (-1,1)\times(-1,1)$, $\mathbf{f}(x_1,x_2) := \operatorname{sign}(x_1)\operatorname{sign}(x_2)\,(1,1)$, $\mathbf{u}^{(0)} := \mathbf{0}$, $p^{(0)} := 0$
3 pre– and 3 post–smoothing steps, $\varepsilon/\varepsilon^{(0)} = \varepsilon_{\text{feas}}/\varepsilon_{\text{feas}}^{(0)} = 10^{-3}$:

level l	$\dim V_l$	$\dim Z_l$	point additive smoother		block multiplicative smoother	
			outer/PCG **iterations**	**total PCG iterations**	**outer**/PCG **iterations**	**total PCG iterations**
1	56	32	**6**/1,0,1,2,4,8	**16**	**6**/1,0,1,2,4,8	**16**
2	208	128	**6**/1,0,1,4,16,30	**52**	**6**/1,0,1,2,5,13	**22**
3	800	512	**5**/1,1,4,20,41	**67**	**6**/1,0,1,2,5,14	**23**
4	3136	2048	**5**/1,1,3,16,47	**68**	**6**/1,0,1,2,6,14	**24**
5	12416	8192	**5**/1,1,3,17,50	**72**	**6**/1,0,1,2,6,15	**25**
6	49408	32768	**5**/1,1,3,19,54	**77**	**6**/1,0,1,2,6,16	**26**

Outline

- Motivation: Large–scale discretized topology optimization
- Setting of the saddle–point problem
- Optimal Uzawa–type methods
- Semi–monotonic augmented Lagrangian methods
- Multigrid preconditioning
- An application to the Stokes problem, numerical results
- Outlook: Applications in topology optimization

Outlook: Applications in Topology Optimization

$$\min_{q \in Q} J(q,u) \quad \text{s.t.} \quad A(q,u) = b \text{ on } V'$$

Nested approach

$$q^* := \arg\min_{q \in Q} J(q, u(q)), \text{ where } A(q, u(q)) = b \text{ on } V'$$

The Hessian is dense!

All–at–once approach

$$\min_{(q,u) \in Q \times V} \max_{\lambda \in V} \{ J(q,u) + \langle \lambda, A(q,u) - b \rangle \}$$

The Hessian is bigger, but sparse and well–structured!

Outlook: Applications in Topology Optimization

Burger & Mühlhuber, 2002: All–at–once parameter identification

$$\min_{q \in L^2(\Omega)} \frac{1}{2}\|\nabla u - \mathbf{B}_{\text{given}}\|^2_{L^2(\Omega)} + \frac{1}{2}\|q\|^2_{L^2(\Omega)} \quad \text{s.t.} \quad -\operatorname{div}(q\nabla u) = f \text{ in } H^1_0(\Omega)'$$

Leads to a sequence of

$$\begin{pmatrix} I & 0 & \text{sym.} \\ 0 & -\triangle & \text{sym.} \\ -\operatorname{div}(\cdot\nabla u^{(k)}) & -\operatorname{div}(q^{(k)}\nabla\cdot) & 0 \end{pmatrix} \begin{pmatrix} \delta q^{(k)} \\ \delta u^{(k)} \\ \delta\lambda^{(k)} \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

$$\begin{aligned} q^{(k+1)} &:= q^{(k)} + \delta q^{(k)} \\ u^{(k+1)} &:= u^{(k)} + \delta u^{(k)} \\ \lambda^{(k+1)} &:= \lambda^{(k)} + \delta\lambda^{(k)} \end{aligned}$$

Outlook: Applications in Topology Optimization

All–at–once topology optimization

$$\min_{q\in L^2(\Omega)} \frac{1}{2}\|\nabla u - \mathbf{B}_{\text{given}}\|^2_{L^2(\Omega)} + \frac{1}{2}\|q\|^2_{L^2(\Omega)} \quad \text{s.t.} \quad -\operatorname{div}(\nu(q)\nabla u) = f \text{ in } H^1_0(\Omega)'$$

Leads to a sequence of

$$\begin{pmatrix} I + \langle \lambda^{(k)}, A''_{qq}(q^{(k)}, u^{(k)})(\cdot)\rangle & \text{sym.} & \text{sym.} \\ \langle \lambda^{(k)}, A''_{uq}(q^{(k)}, u^{(k)})(\cdot)\rangle & -\triangle + \langle \lambda^{(k)}, A''_{uu}(q^{(k)}, u^{(k)})(\cdot)\rangle & \text{sym.} \\ -\operatorname{div}(\nu'_q(q^{(k)})(\cdot)\nabla u^{(k)}) & -\operatorname{div}(\nu(q^{(k)})\nabla\cdot) & 0 \end{pmatrix} \begin{pmatrix} \delta q^{(k)} \\ \delta u^{(k)} \\ \delta\lambda^{(k)} \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

$$\begin{aligned} q^{(k+1)} &:= q^{(k)} + \delta q^{(k)} \\ u^{(k+1)} &:= u^{(k)} + \delta u^{(k)} \\ \lambda^{(k+1)} &:= \lambda^{(k)} + \delta\lambda^{(k)} \end{aligned}$$