A Newton-Multigrid Method for PDE Constrained Optimization

Martin Engel

Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn

Workshop on PDE Constrained Optimization
University of Hamburg
March 27-29, 2008
1. Introduction

2. Newton-Multigrid
   - Constraint Preconditioning
   - Multigrid
   - Numerical Results

3. Control-Constrained Problems
   - Primal-Dual Active Set Strategy
   - Numerical Results

4. Conclusions
Outline

1 Introduction

2 Newton-Multigrid
   - Constraint Preconditioning
   - Multigrid
   - Numerical Results

3 Control-Constrained Problems
   - Primal-Dual Active Set Strategy
   - Numerical Results

4 Conclusions
Problem Formulation

Optimal Control Problem

Minimize \[ J(y, u) = \frac{1}{2} \| y - \bar{y} \|_{L^2}^2 + \frac{\sigma}{2} \| u \|_{L^2}^2 \]
subject to \[ C(y, u) = Ly + N(y) - u - f = 0 \]

- \( Y (= H^1_0(\Omega)) \) and \( U (= L^2(\Omega)) \) Hilbert spaces
- (Distributed) control \( u \in U \)
  (later: \( u \in U_{ad} \subset U \) a closed and convex set)
- Target state \( \bar{y} \in U \)
- \( L \) second-order elliptic operator, \( C : Y \times U \to W \)
Existence and Uniqueness

Lagrangian

\[ L(y, u, p) = J(y, u) - \langle C(y, u), p \rangle_{W, W'} \]

First-Order Optimality Conditions (Equality-Constrained Case)

\[ \nabla L(y^*, u^*, p^*) = \begin{bmatrix} \nabla_y J(y^*, u^*) + (C_y(y^*, u^*))^* p^* \\ \nabla_u J(y^*, u^*) + (C_u(y^*, u^*))^* p^* \\ C(y^*, u^*) \end{bmatrix} = 0 \]
Outline

1 Introduction

2 Newton-Multigrid
   - Constraint Preconditioning
   - Multigrid
   - Numerical Results

3 Control-Constrained Problems
   - Primal-Dual Active Set Strategy
   - Numerical Results

4 Conclusions
Applying Newton’s method to the first-order conditions yields a sequence of quadratic problems (QPs):

**Newton Step**

\[
\begin{bmatrix}
L_{yy}(y, u, p) & L_{yu}(y, u, p) & (C_y(y, u))^* \\
L_{uy}(y, u, p) & L_{uu}(y, u, p) & (C_u(y, u))^* \\
C_y(y, u) & C_u(y, u) & 0
\end{bmatrix}
\begin{bmatrix}
\delta y \\
\delta u \\
\delta p
\end{bmatrix}
= -
\begin{bmatrix}
\nabla_y J(y, u) + (C_y(y, u))^* p \\
\nabla_u J(y, u) + (C_u(y, u))^* p \\
C(y, u)
\end{bmatrix}
\]
The Discrete QPs

The (discretized) QP in each Newton step is a linear Saddle Point System

\[
\begin{bmatrix}
W_{11} & W_{12} & C_1^T \\
W_{21} & W_{22} & C_2^T \\
C_1 & C_2 & 0
\end{bmatrix}
\begin{bmatrix}
y \\
u \\
p
\end{bmatrix}
=
\begin{bmatrix}
f \\
g \\
h
\end{bmatrix}
\]  

Solution Methods

- Direct factorization methods (memory, complexity)
- Iterative methods (efficient preconditioning)
- Multigrid (suitable relaxation method)
Full Space vs. Reduced Space Methods

**Reduced Space**

- Elimination yields reduced system in (lower dimensional) space of controls
- Feasibility in each step
- Can increase nonlinearity ([Ghattas and Bark])

**Full Space**

- Simultaneous solution for all unknowns
- Feasibility at minimizer
- Ill-conditioned system requires effective preconditioning

Remark: preconditioning, relaxation and different degrees of inexactness tend to blur the line between reduced and full space methods.
Multigrid in Optimization

Multigrid in Inner Iterations

- MG for Fredholm operator of the 2nd kind [Hackbusch '80]
- MG solver for state and adjoint in RSQP [Schulz '96]

MG for Outer Iteration

- Collective Relaxation [Brandt '84, Vanka '86, Borzi '02, Ascher and Haber '03]
- Relaxation as inexact block-factorization
  - Gradient Descent [Arian and Ta’asan '94]
  - Range Space Iterations [Braess and Sarazin '97, Wittum '89]
  - Null Space Iterations [Maar and Schulz '00]
  - Constraint Preconditioner (based on [Keller, Gould and Wathen '00])
- MG/Opt [Brandt '84, Lewis and Nash '00, Borzi '05]
Consider a *constraint preconditioner*

\[
P = \begin{pmatrix}
G_{11} & G_{12} & C_1^T \\
G_{21} & G_{22} & C_2^T \\
C_1 & C_2 & 0
\end{pmatrix}
\quad \text{for} \quad
K = \begin{pmatrix}
W_{11} & W_{12} & C_1^T \\
W_{21} & W_{22} & C_2^T \\
C_1 & C_2 & 0
\end{pmatrix}
\]

Then the preconditioned matrix \( P^{-1}K \) has ([Keller, Gould and Wathen ’00]):

- an eigenvalue 1 with multiplicity 2N
- \( M \) eigenvalues which are defined by the generalized eigenvalue problem

\[
Z^T WZx = \lambda Z^T GZx,
\]

where \( Z \) is a basis for the nullspace of the linearized constraints
- the dimension of \( \mathcal{K}(P) \) is at most \( M+2 \)
Constraint Preconditioner, cont’d

\[ P_{CP} = \begin{pmatrix} 0 & 0 & C_1^T \\ 0 & B_Z & C_2^T \\ C_1 & C_2 & 0 \end{pmatrix} \]

\(P_{CP}\)-GMRES converges in three iterations, provided that the true reduced Hessian is used, i.e.

\[ B_Z = H_Z = Z^T WZ, \]

where

\[ H_Z = C_2^T C_1^{-T} W_{11} C_1^{-1} C_2 - C_2^T C_1^{-T} W_{12} - W_{21} C_1^{-1} C_2 + W_{22} \]

for the fundamental basis \( Z = \begin{bmatrix} -C_2^T C_1^{-T} & I \end{bmatrix} \)

- 2 PDE solves \((C_1^{-1}, C_1^{-T})\)
- 2 PDE solves per MatVec with \(H_Z\)
Inexact Constraint Preconditioner as Smoother

One application of $P^{-1}_{CP}$ results in the following algorithm:

1. solve for $p$: $C_1^T p = f_y$
2. solve for $u$: $Hzg = f_u - C_2^T p$
3. solve for $y$: $C_1 y = f_p - C_2 u$

Inexact version of $\tilde{P}_{CP}$:

- replace the solve in 1. with relaxation $\tilde{C}_1^T$ for $C_1^T$
- replace the solve in 3. with relaxation $\tilde{C}_1$ for $C_1$
- solve 2. with a fixed (small) number of CG iterations.
- use relaxations $\tilde{C}_1^T$ and $\tilde{C}_1$ to approximate matrix-vector product with $Hz$ in 2

Use $\tilde{P}_{CP}$ as preconditioner in Richardson iteration
Illustration of Smoothing Behaviour

Error in controls, initial value and iterations 2,3,4.
\[-\nabla \cdot K \nabla y = f \quad \text{in } \Omega\]

\[y = 0 \quad \text{on } \partial \Omega\]

\(F : \hat{\Omega} \rightarrow \Omega\) smooth map, \(RT_0\)-spaces are reduced to cell-centered FD by applying suitable quadrature rules [Arbogast, Wheeler, Yotov]

- **Standard coarsening:** \(h_k = h_{\text{coarse}} 2^{-k}\)
- **Direct coarse grid discretization.** For \(h_k, k = 0, \ldots, L - 1\)
  - Grid sequence \(\hat{\Omega}_{h_k}\)
  - Discretized constraints \(L_{h_k} y_{h_k} = M_{h_k} u_{h_k}\), i.e. on level \(k\) we have \(C_1 = L_{h_k}\) and \(C_2 = M_{h_k}\)
Multigrid Components II

- Coarse grid solver is 3-step constraint-preconditioned Krylov method
- V-, W-, and F-Cycles with $\nu_1$ pre- and $\nu_2$ postsmoothing steps

$$x_k^{m+1} \leftarrow MG^\gamma(k, b_k, x_k^m) \quad (\gamma = 1 : V, \gamma = 2 : W)$$

if $k = 0$ then
  Solve $K_0x_0 = b_0$
else
  Presmooth $\tilde{x}_k = S^{\nu_1}x_k^m$
  Residual restriction $b_{k-1} = I_k^{k-1}(b_k - K_k\tilde{x}_k)$
  Grid recursion $v_{k-1} = MG^{\gamma}(k - 1, b_{k-1}, 0)$
  Correction $\tilde{x}_k = \tilde{x}_k + I_{k-1}^k v_{k-1}$
  Postsmooth $x_k^{m+1} = S^{\nu_2}\tilde{x}_k$
end if

- Variable V-Cycle: $\nu_k = 2^{L-1-k}\nu_f$
- F-Cycle recursively defined as F-Cycle followed by V-Cycle
Interpolation Operators for CCFD

- Blockwise definition

\[ I_h^H = \begin{pmatrix} I_h^H & I_h^H \\ I_h^H & I_h^H \end{pmatrix} \]

- \( I_h^H \) four point average (FPA)

\[ I_{h,FPA}^H = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \]

- \( I_h^H \) bilinear (BL) or Wesseling/Khalil (WK) interpolation

\[ I_{h,\text{BL}}^H = \frac{1}{16} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 3 & 9 & 9 & 3 \\ 3 & 9 & 9 & 3 \\ 1 & 3 & 3 & 1 \end{bmatrix}, \quad I_{h,\text{WK}}^H = \frac{1}{16} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \]

- \( m_p + m_r = 4 > 2 \)
## Numerical Results, MG, Linear Model Problem

### Avg. rates of convergence for different cycle types on the Linear Model Problem

<table>
<thead>
<tr>
<th>Cycle \</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>V(1,1)</td>
<td>1.02_1</td>
<td>1.08_1</td>
<td>1.13_1</td>
<td>1.14_1</td>
<td>1.14_1</td>
<td>1.14_1</td>
<td>1.15_1</td>
<td>1.15_1</td>
</tr>
<tr>
<td>V(2,1)</td>
<td>6.37_2</td>
<td>6.70_2</td>
<td>7.07_2</td>
<td>7.21_2</td>
<td>7.30_2</td>
<td>7.31_2</td>
<td>7.31_2</td>
<td>7.31_2</td>
</tr>
<tr>
<td>V(2,2)</td>
<td>4.78_2</td>
<td>4.85_2</td>
<td>5.15_2</td>
<td>5.26_2</td>
<td>5.34_2</td>
<td>5.35_2</td>
<td>5.35_2</td>
<td>5.35_2</td>
</tr>
<tr>
<td>F(1,1)</td>
<td>8.18_2</td>
<td>8.16_2</td>
<td>8.18_2</td>
<td>8.20_2</td>
<td>8.20_2</td>
<td>8.20_2</td>
<td>8.21_2</td>
<td>8.21_2</td>
</tr>
<tr>
<td>W(1,1)</td>
<td>8.18_2</td>
<td>8.16_2</td>
<td>8.18_2</td>
<td>8.20_2</td>
<td>8.20_2</td>
<td>8.20_2</td>
<td>8.21_2</td>
<td>8.21_2</td>
</tr>
</tbody>
</table>

![Graph showing L_2 norm of error vs. wall clock time in seconds]
The Full Multigrid Algorithm

Full Multigrid

\[ k = 0: \]
\[ x_0^{\text{FMG}} \leftarrow Kx_0 = b_0 \]
\[ \text{for } k = 1, 2, \ldots, L \text{ do} \]
\[ x_k^0 = \prod_{k-1}^k x_{k-1}^{\text{FMG}} \]
\[ x_k^{\text{FMG}} = MG^r(x_k^0, k + 1, \gamma) \]
\[ \text{end for} \]

FMG Properties

- $\mathcal{O}(N)$ operations to compute $x_h^{\text{FMG}}$ ($W_l^{\text{FMG}} = \frac{4}{3} rW_l$ in 2D)
- $\|x_l - x_l^{\text{FMG}}\| \sim \|x - x_l\|$ ($\|x_l - x_l^{\text{FMG}}\| \leq C_{\rho, r} h^\alpha$, $\|M_l\| \leq \rho < 1$)
Numerical Results, FMG, Linear Model Problem

$L_2$-norm of discretization error and wall-clock time in seconds for one FMG cycle. On each level, a $V(1,1)$ cycle is used.

<table>
<thead>
<tr>
<th>Level</th>
<th>$h$</th>
<th>$N$</th>
<th>Time[s]</th>
<th>ratio</th>
<th>$L_2$-Error</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1/128</td>
<td>49152</td>
<td>0.1680</td>
<td>—</td>
<td>6.06389E-05</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>1/256</td>
<td>196608</td>
<td>0.7578</td>
<td>4.511</td>
<td>1.59350E-05</td>
<td>0.263</td>
</tr>
<tr>
<td>9</td>
<td>1/512</td>
<td>786432</td>
<td>3.6797</td>
<td>4.856</td>
<td>4.06443E-06</td>
<td>0.255</td>
</tr>
<tr>
<td>10</td>
<td>1/1024</td>
<td>3145728</td>
<td>16.4102</td>
<td>4.459</td>
<td>1.02422E-06</td>
<td>0.252</td>
</tr>
<tr>
<td>11</td>
<td>1/2048</td>
<td>12582912</td>
<td>68.3516</td>
<td>4.165</td>
<td>2.56857E-07</td>
<td>0.251</td>
</tr>
<tr>
<td>12</td>
<td>1/4096</td>
<td>50331648</td>
<td>276.5470</td>
<td>4.041</td>
<td>6.42940E-08</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Costs of Relaxation

- MatVec $K$: $2C_N$
- Constraint blocks: $2C_N$
- MatVec $B_Z$ (CG iter): $2C_N$
- MatVec $B_Z$ (CG init): $2C_N$

Total: $8C_N$, plus: inner products, ...
Convergence and $L^2$-Regularization

For small $\sigma$, the coarse grid correction with discrete $H_Z$ is not effective

- Use more robust cycle (F, W, variable-V)
- Use MG as preconditioner for GMRES
- Preconditioning $H_Z$?
- Learn from Helmholtz/Convection-Diffusion?
- Choosing appropriate coarse grid size $h_c$ only “sure thing”
Numerical Results, Diffusion Problem

\[
K = \begin{bmatrix}
11 & 9 \\
9 & 13
\end{bmatrix}
\]

Target state $\bar{y}$ and controls $u$ (top left to right), $L_2$-error of $y$ and KKT residual.

ALGS-Smoother for state and adjoint.

Avg. residual reduction rate 0.0987.
Numerical Results, Non-Uniform Grid

Mesh deformation (top left), convergence history on Level 11 for \( \delta = 0.1 \) (top right), \( \delta = 0.15 \) (bottom left), \( \delta = 0.2 \) (bottom right).

Table: convergence rates for \( \delta = 0.2 \).

PDE relaxation with ALGS (similar results are obtained for ILU smoothing.)

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{2,2} ) PDE</td>
<td>0.0390</td>
<td>0.0502</td>
<td>0.0571</td>
<td>0.0622</td>
<td>0.0654</td>
</tr>
<tr>
<td>( V_{1,1} ) PDE</td>
<td>0.0997</td>
<td>0.1155</td>
<td>0.1217</td>
<td>0.1301</td>
<td>0.1313</td>
</tr>
<tr>
<td>( V_{1,1} ) KKT</td>
<td>0.1332</td>
<td>0.1352</td>
<td>0.1363</td>
<td>0.1350</td>
<td>0.1316</td>
</tr>
</tbody>
</table>
Inexact Newton Method

Fully converging iterative solvers is neither reasonable (oversolving) nor feasible (too costly), thus solve the Newton system

$$\nabla F(x_k) \Delta x_k = -F(x)$$

to some tolerance → inexact Newton method.

Stopping Criterion

$$\| r_k^{(i)} \| \leq \eta_k \| r_k \|$$

Choosing the forcing sequence $\eta_k$ as

- $\eta_k \leq \eta < 1$ guarantees convergence
- $\eta_k \to 0$ yields superlinear convergence
- $\eta_k = O(\| r_k \|)$ yields quadratic convergence

E.g., choose $\eta_k = \min(c \| F(x_k) \|^p, 0.5)$, $0 < p \leq 1$
Semilinear Problem

\[ C(y, u) = -\Delta y + \gamma ye^y - u. \] Convergence history for \( \gamma = 1 \) (left) and \( \gamma = 10 \).
Outline

1 Introduction

2 Newton-Multigrid
   - Constraint Preconditioning
   - Multigrid
   - Numerical Results

3 Control-Constrained Problems
   - Primal-Dual Active Set Strategy
   - Numerical Results

4 Conclusions
Control Constraints

- Unilateral constraints on the control

\[ u \in U_{ad} = \{ v \in U | v \leq u_{bnd} \text{ a.e. in } \Omega \} \]

- Variational inequality

\[ \hat{J}(u^*)(u - u^*) \geq 0 \text{ for all } u \in U_{ad} \]

- Using \( \hat{J}(u) = H_Z u - L^{-1} \bar{y} \) and substituting \( p \) yields

\[ (\sigma u^* - p^*, u - u^*) \geq 0 \text{ for all } u \in U_{ad} \]

- Complimentarity condition

\[ \sigma u^* - p^* + \lambda = 0, \quad \lambda(u^* - u_{bnd}) = 0, \quad \lambda \geq 0 \text{ a.e.} \]
Primal Dual Active Set Strategy

The equivalent form $\lambda = c \max(0, u^* + \frac{\lambda}{c} - u_{\text{bnd}})$, $c > 0$ for the complementarity condition gives rise to

Primal-Dual Active Set Method

\begin{itemize}
  \item $k = 0$, choose $y^0, u^0, p^0, \lambda^0$
  \item while Not converged do
    \begin{itemize}
      \item Predict $\mathcal{A}_k, \mathcal{I}_k$
      \item if $n \geq 2$ and $\mathcal{A}_k = \mathcal{A}_{k+1}$ then Converged
      \item else
        \begin{itemize}
          \item Solve EQP
        \end{itemize}
      \end{itemize}
    \end{itemize}
  \item $k = k + 1$
\end{itemize}

Active Set Prediction

\begin{align*}
  \mathcal{A}_k &= \{ x | u_{k-1} + \frac{\lambda_{k-1}}{\sigma} > u_{\text{bnd}} \} \\
  \mathcal{I}_k &= \Omega \setminus \mathcal{A}_k
\end{align*}

[Hintermüller, Ito, Kunisch]
The PDAS EQP

The EQP (semismooth Newton step) is given by

**PDAS System**

\[
\begin{align*}
M y^{k+1} + L^T p^{k+1} &= \tilde{y} \\
\sigma M u^{k+1} + M^T p^{k+1} + \lambda^{k+1} &= 0 \\
L y^{k+1} + M u^{k+1} &= 0 \\
\lambda^{k+1} &= 0 \text{ on } \mathcal{I}_k \\
u^{k+1} &= u_{\text{bnd}} \text{ on } \mathcal{A}_k
\end{align*}
\]

- Use splitting \( u^k = [u_{\mathcal{I}_k} \ u_{\mathcal{A}_k}] \) to reduce EQP to a system for \( y^k, u_{\mathcal{I}_k}, p^k \)
- Solve rEQP with our Multigrid method
- \( B_{\mathcal{Z},\mathcal{I}} = \sigma M_{\mathcal{I},\mathcal{I}} + M_{\mathcal{I}} L^{-T} M L^{-1} M_{\mathcal{I}} \)
- Set \( \lambda_{\mathcal{A}}^{k+1} = R_{\mathcal{A}}(\sigma M_{\mathcal{I}} u_{\mathcal{I}}^{k+1} + M p^{k+1}) \)
Restriction of Active Sets

Active Sets on coarse grids are constructed by piecewise constant restriction of $\chi_A$ with the condition $A_{2h} \subset A_h$ and $I_{2h} \supset I_h$ for each $h$.

Active (red) and inactive (blue) sets for $h = 2^{-k}$, $k = 6, 5, 4, 3$. 
Control-Constrained Model Problems

Computed optimal controls $u$ on $h = 2^{-6}$ mesh for different upper bound functions $u_b$.
Numerical Results for PDAS

### PDAS Iteration $L^2$-Error of Control $u$

<table>
<thead>
<tr>
<th>Level</th>
<th>$e_{L^2}$</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>$2.9225 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$7.3051 \times 10^{-6}$</td>
<td>$4.0005$</td>
</tr>
<tr>
<td>8</td>
<td>$1.8262 \times 10^{-6}$</td>
<td>$4.0001$</td>
</tr>
<tr>
<td>9</td>
<td>$4.5655 \times 10^{-7}$</td>
<td>$4.0000$</td>
</tr>
<tr>
<td>10</td>
<td>$1.1414 \times 10^{-7}$</td>
<td>$4.0000$</td>
</tr>
<tr>
<td>11</td>
<td>$2.8534 \times 10^{-8}$</td>
<td>$4.0000$</td>
</tr>
</tbody>
</table>

### PDAS Iteration $L^2$-Norm of $e_{bnd}$

### Multigrid Iterations Residual Norm
### Numerical Results II

<table>
<thead>
<tr>
<th>σ</th>
<th>L</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>124952</td>
<td>+6096</td>
<td>+24</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>6.843_{-2}</td>
<td>4.149_{-4}</td>
<td>0.0</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1.0_{-2}</td>
<td>10</td>
<td>499670</td>
<td>+24474</td>
<td>+144</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>6.848_{-2}</td>
<td>4.800_{-4}</td>
<td>0.0</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>1998192</td>
<td>+98254</td>
<td>+706</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>6.856_{-2}</td>
<td>4.820_{-4}</td>
<td>0.0</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>101450</td>
<td>+20528</td>
<td>+7744</td>
<td>+1330</td>
<td>+20</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2.338_{-1}</td>
<td>9.117_{-2}</td>
<td>8.090_{-3}</td>
<td>3.379_{-4}</td>
<td>0.0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>1.667</td>
<td>9.732_{-1}</td>
<td>5.163</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>1622360</td>
<td>+329360</td>
<td>+123534</td>
<td>+21352</td>
<td>+546</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2.331_{-1}</td>
<td>9.109_{-2}</td>
<td>1.055_{-2}</td>
<td>3.960_{-4}</td>
<td>0.0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>1.675</td>
<td>1.271</td>
<td>3.555</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Size of active sets, $\|u_h - u_{bnd,h}\|$ and ratio $e_k/e_{k-1}$ (for $\sigma = 1.0_{-5}$ only) for levels 9, 10, 11 and PDAS iteration $k$. 

A Newton-Multigrid Method for PDE-Constrained Optimization
Martin Engel PDE-Opt Workshop ’08
Numerical Results III

![Graphs showing numerical results for different PDAS Iterations and L^2 Error of u.](image)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
</table>
| bnd 1 | A = 1998192 | 98254 | 706
|       | \(e_{bnd} = 6.856_{-2} \) | 4.820_{-4} | 0.0 |
| bnd 2 | A = 2536106 | 106944 | 742
|       | \(e_{bnd} = 9.4103_{-2} \) | 7.106_{-4} | 0.0 |
| bnd 3 | A = 2684033 | 128150 | 808
|       | \(e_{bnd} = 1.081_{-1} \) | 7.547_{-4} | 0.0 |
| bnd 4 | A = 2856365 | 123806 | 797
|       | \(e_{bnd} = 1.305_{-1} \) | 8.618_{-4} | 0.0 |
Numerical Results IV

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2044753</td>
<td>+703069</td>
<td>+104690</td>
<td>+12471</td>
<td>+877</td>
<td>+80</td>
<td>+2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>2007484</td>
<td>+185614</td>
<td>+3921</td>
<td>+27</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>$\varepsilon = 10^{-10}$</td>
<td>1998192</td>
<td>+98254</td>
<td>+706</td>
<td></td>
<td></td>
<td></td>
<td>32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>10</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
One FMG cycle per PDAS step recovers mesh-independent behaviour.

<table>
<thead>
<tr>
<th>Level</th>
<th>$e_{L^2}$</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>$8.36991 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$2.12820 \times 10^{-5}$</td>
<td>3.9329</td>
</tr>
<tr>
<td>9</td>
<td>$5.35098 \times 10^{-6}$</td>
<td>3.9772</td>
</tr>
<tr>
<td>10</td>
<td>$1.34025 \times 10^{-7}$</td>
<td>3.9925</td>
</tr>
<tr>
<td>11</td>
<td>$3.35271 \times 10^{-7}$</td>
<td>3.9975</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7794</td>
<td>+398</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>31158</td>
<td>+1604</td>
<td>+6</td>
</tr>
<tr>
<td>9</td>
<td>124952</td>
<td>+6092</td>
<td>+28</td>
</tr>
<tr>
<td>10</td>
<td>499670</td>
<td>+24474</td>
<td>+144</td>
</tr>
<tr>
<td>11</td>
<td>1998192</td>
<td>+98242</td>
<td>+718</td>
</tr>
</tbody>
</table>
Conclusions

Summary:
- Multigrid for optimal control problems
- Control-constraints via PDAS
- Full multigrid solves optimal control problems in $O(n)$
- State/adjoint-specific smoothers (ALGS, ILU)

Future work:
- Extension to 3D
- Parallelization
References I

E. Arian and S. Ta’asan.  
Multigrid one-shot methods for optimal control problems: infinite dimensional control.  

U. M. Ascher and E. Haber.  
A multigrid method for distributed parameter estimation problems.  

A. Borzi and K. Kunisch.  
A multigrid scheme for elliptic constrained optimal control problems.  

D. Braess and R. Sarazin.  
An efficient smoother for the Stokes problem.  
References II

W. Hackbusch.
Fast solution of elliptic control problems.

Constraint preconditioning for indefinite linear systems.

A multigrid approach to the optimization of systems governed by differential equations.
