Multilevel Optimization Techniques Using the Exact Hessian for PDAE Constrained Optimization Problems

TECHNISCHE UNIVERSITÄT DARMSTADT

GAMM
Trier, 2009

Debora Clever

joint project with C. Ziems, S. Ulbrich and J. Lang

SPP 1253, supported by the German Research Foundation (DFG)

Technische Universität Darmstadt

Outline

Multilevel Second Order Optimization
Motivation

- $\min_u \hat{J}(u) := J(y,u)$ s.t. $e(y,u) = 0$
- PDAE constrained optimization requires several solves of complex time-dependent differential equations

$\Rightarrow$ Design efficient second order optimization algorithms which
  - reduce the number of PDAE solves
  - carry out as many solves on coarse grids as possible but still reasonable
  - cooperate with fully adaptive state of the art PDAE solvers, like KARDOS

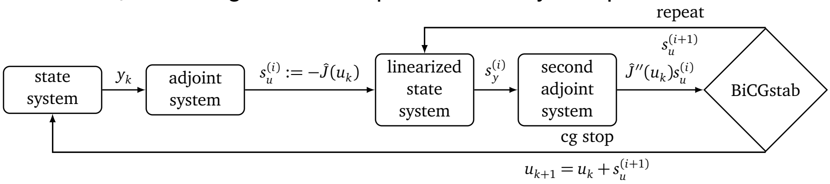
- We use:
  - Derivative based optimization technique
  - Backwards approach
  - “First optimize then discretize” strategy

Multilevel Second Order Optimization
Control Update

- Control update s_u with $u_{k+1} = u_k + s_u$ is computed by solving

$$\hat{J}''(u_k)s_u = -\hat{J}'(u_k)$$

  with reduced gradient $\hat{J}'(u_k)$ and reduced Hessian $\hat{J}''(u_k)$
- Reduced Hessian can not be computed on its own but only applied to a direction s_u which again is the input of this very computation



Multilevel Second Order Optimization
Control Update

- Control update s_u with $u_{k+1} = u_k + s_u$ is computed by solving

$$\hat{J}''(u_k)s_u = -\hat{J}'(u_k)$$

  with reduced gradient $\hat{J}'(u_k)$ and reduced Hessian $\hat{J}''(u_k)$
- Reduced Hessian can not be computed on its own but only applied to a direction s_u which again is the input of this very computation

$\ominus$ Control update s_u must be computed within appropriate iterative method

$\oplus$ Control update s_u carries direction and step length

$\oplus$ Superlinear or even quadratic convergence

Multilevel Second Order Optimization
Grid Adjustment

Q When to stop inner iteration?

Multilevel Second Order Optimization
Grid Adjustment

TECHNISCHE UNIVERSITÄT DARMSTADT

Q When to stop inner iteration?

A In accordance to accuracy of outer iteration $\Rightarrow$ $\text{tol}_i = \text{tol}_o^{1.5}$. (superlinear convergence)

Multilevel Second Order Optimization
Grid Adjustment

TECHNISCHE UNIVERSITÄT DARMSTADT

Q When to stop inner iteration?

A In accordance to accuracy of outer iteration $\Rightarrow$ $\text{tol}_i = \text{tol}_o^{1.5}$. (superlinear convergence)

Q How to choose tolerances for outer iteration?

Multilevel Second Order Optimization
Grid Adjustment

TECHNISCHE UNIVERSITÄT DARMSTADT

Q When to stop inner iteration?

A In accordance to accuracy of outer iteration $\Rightarrow$ $\text{tol}_i = \text{tol}_o^{1.5}$. (superlinear convergence)

Q How to choose tolerances for outer iteration?

A In accordance to distance of current control and optimal control

Multilevel Second Order Optimization
Grid Adjustment

Q When to stop inner iteration?

A In accordance to accuracy of outer iteration $\Rightarrow \; \text{tol}_i = \; \text{tol}_o^{1.5}$. (superlinear convergence)

Q How to choose tolerances for outer iteration?

A In accordance to distance of current control and optimal control

Q How to measure this distance?

Multilevel Second Order Optimization
Grid Adjustment

TECHNISCHE UNIVERSITÄT DARMSTADT

Q When to stop inner iteration?

A In accordance to accuracy of outer iteration $\Rightarrow$ $\text{tol}_i = \text{tol}_o^{1.5}$. (superlinear convergence)

Q How to choose tolerances for outer iteration?

A In accordance to distance of current control and optimal control

Q How to measure this distance?

A The reduced gradient descents towards zero as approaching the optimum

Refine if global state and adjoint errors in space and time become greater then a multiple of reduced gradient

If point wise control constraints are present consider projected gradient

Glass Cooling
Motivation

- Glass is processed at temperatures up to 1000°C
- Expeditious cooling at room temperature causes cracks inside the material

- Solution: Controlled cooling within a preheated furnace by decreasing its temperature from 700-800°C to 20°C
- Radiation plays a dominant role because of high temperatures

$\Rightarrow$ Optimization of furnace temperature results in an optimal boundary control problem with high dimensional and highly nonlinear partial differential algebraic equations in at least two components

Glass Cooling Modelling

Dimensionless state system with glass temperature $T(x,t)$, mean radiation intensity $\phi(x,t)$ and furnace temperature $u(t)$ (control)

$$
\begin{aligned}
\partial_t T - k\Delta T - \frac{1}{3\kappa}\Delta\phi &= 0 && \text{in } \Omega \times (0, t_e] \\
-\frac{\epsilon^2}{3\kappa}\Delta\phi &= -\kappa\phi + 4\pi\kappa a T^4 && \text{in } \Omega \times (0, t_e] \\
kn\cdot\nabla T + \frac{1}{3\kappa} n\cdot\nabla\phi &= \frac{h}{\epsilon}(u - T) + \frac{1}{2\epsilon}4\pi a u^4 - \frac{1}{2\epsilon}\phi && \text{in } \partial\Omega \times (0, t_e] \\
\frac{\epsilon^2}{3\kappa} n\cdot\nabla\phi &= \frac{\epsilon}{2}(4\pi a u^4 - \phi) && \text{in } \partial\Omega \times (0, t_e] \\
T(0,x) &= T_0(x) && \text{on } \Omega.
\end{aligned}
$$

Gray scale model, A. Klar, R. Pinnau, G. Thömmes et al.

Optimal Control Problem
Objective Functional

TECHNISCHE UNIVERSITÄT DARMSTADT

- Temperature T should follow desired profile T_d
- Regularization of furnace temperatue $u(t)$

$$J(T,u) := \frac{1}{2}\int_0^{t_e} \|T - T_d\|^2_{L^2(\Omega)}\mathrm{d}t + \frac{\delta_e}{2}\|(T - T_d)(t_e)\|^2_{L^2(\Omega)} + \\ + \frac{\delta_u}{2}\int_0^{t_e} \|u - u_d\|^2_{L^2(\partial\Omega)}\mathrm{d}t$$

- The separate tracking of the final state is important to obtain non-vanishing gradient or Hessian information in the last time step

Cooperation with M. Herty, RWTH Aachen

Numerical Experiments
Setting

- Present performace of multilevel strategy
- Give comparative values (disable multilevel strategy, gradient method)

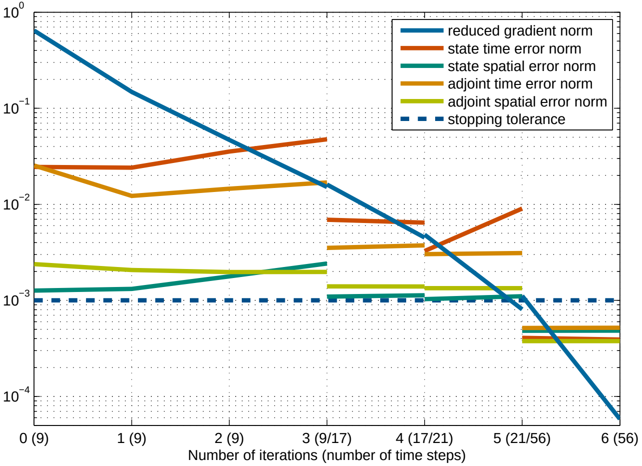
- Computational domain $[-2, 2] \times [-2, 2] \in \mathbb{R}^2$, time $t \in [0, 0.1]$
- $T(x, 0) = 900$, u(t) spatially constant
- $u_0 = u_d = T_d$: exponential decrease from 900 to 400
- Set $\delta_u = 0.1$, $\delta_e = 0.1$

Numerical Experiments
Multilevel SQP-Method, Glass Cooling

TECHNISCHE UNIVERSITÄT DARMSTADT

- Start with 9 time steps and 144 spatial nodes
- Grids are automatically adjusted if global error estimators exceed reduced gradient norm
- Algorithm is stopped if reduced gradient norm and all error estimators are less than 1.0e-3
- Optimal control is computed with 56 time points and up to 4239 spatial nodes

Numerical Experiments
Multilevel Strategy



Numerical Experiments
Optimization Protocol

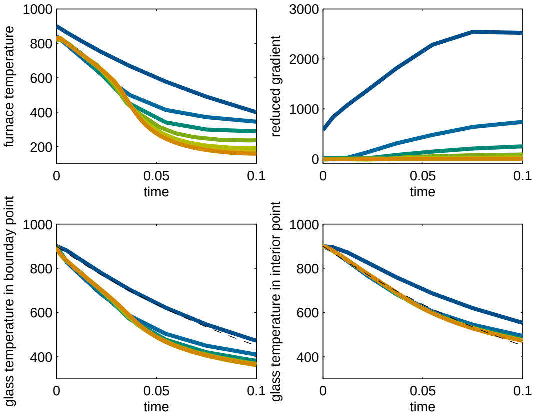
TECHNISCHE UNIVERSITÄT DARMSTADT

it	objective	reduced grad. norm	maximal error norm	time tolerance	#time steps	#space nodes	#cg iter	reason to stop bicg	CPU
1	1.8163e+04	6.4434e-01	2.5392e-02	5.00e-02	9	144			3.83
2	4.4280e+03	1.4790e-01	2.4077e-02	5.00e-02	9	144	2	res small	15.67
3	2.8917e+03	4.6834e-02	3.5476e-02	5.00e-02	9	144	2	res small	27.66
4	2.5099e+03	1.5117e-02	4.7599e-02	5.00e-02	9	144	3	res small	43.97
4	2.4997e+03	1.6096e-02	6.9082e-03	1.59e-02	17	144			48.29
5	2.4038e+03	4.5038e-03	6.4502e-03	1.59e-02	17	144	4	max iter	77.31
5	2.4053e+03	4.8374e-03	3.2651e-03	8.24e-03	21	144			85.58
6	2.3925e+03	8.0697e-04	9.0239e-03	8.24e-03	21	144	4	max iter	119.55
6	2.3578e+03	1.1139e-03	5.1654e-04	7.36e-04	56	144-4239			148.95
7	2.3570e+03	5.8087e-05	5.1826e-04	7.36e-04	56	144-4239	4	max iter	473.02

- Four optimization steps on coarse grid with only 2-3 BiCGstab iterations
- Only one optimization step on each refined grid

Numerical Experiments
Multilevel SQP-Method, Glass Cooling

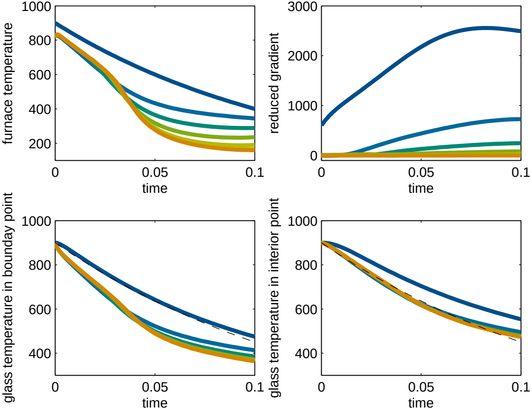
TECHNISCHE UNIVERSITÄT DARMSTADT



- SQP with multilevel strategy → CPU: 473 seconds (8 minutes)

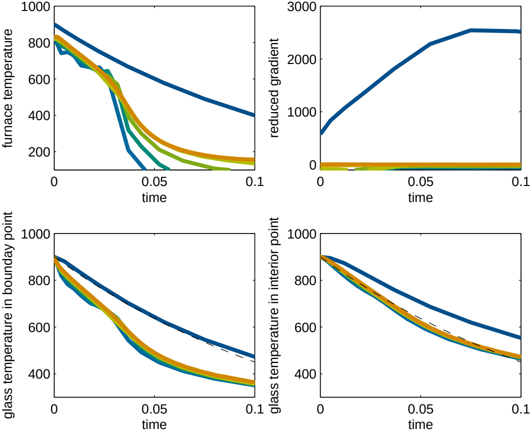
Numerical Experiments
Improvement: Multilevel Strategy

TECHNISCHE UNIVERSITÄT DARMSTADT



- SQP without multilevel → CPU: 4073 seconds (1 hour and 8 minutes)

Numerical Experiments
Improvement: Second Order Derivative



- Gradient method with multilevel → CPU: 2853 seconds (48 minutes)
- Gradient method without multilevel → CPU: $\approx$ 25000 seconds (7 hours)

Numerical Experiments
Multilevel SQP-Method, Glass Cooling

- Second order information reduces CPU by a factor six
- Multilevel strategy reduces CPU by a factor ten
- Overall saving: factor 60

  (which means a saving of 98% in comparison to the fixed gradient method)

- All final results are comparable

Multilevel Optimization Techniques
Summary and Current Research

TECHNISCHE UNIVERSITÄT DARMSTADT

Summary

- Derivation of a multilevel SQP-method using the exact reduced Hessian (in cooperation with C.Ziems, S. Ulbrich, TU Darmstadt)
- Final state tracking to improve derivative information (in cooperation with M. Herty, RWTH Aachen)
- Application to glass cooling problem, modelled by radiative heat transfer (gray scale model, A. Klar, R. Pinnau, G. Thömmes et al.)
- CPU reduction by approximately 98% in comparison to a gradient method with simple line search on a fixed grid

Current research

- Augment objective functional with glass temperature gradient and corresponding final state term ($\rightarrow$ theory)
- Application to N-band modell to take frequency dependent behavior of radiation into account
- Consideration of control constraints as additional minimization problem to ensure proper step size