Efficient numerical solution of large scale LQR problems arising in the optimal control of parabolic PDEs

Jens Saak
joint work with Peter Benner, Sabine Görner
and Hermann Mena (EPN Quito Ecuador)

Professur Mathematik in Industrie und Technik
Fakultät für Mathematik
Technische Universität Chemnitz

February 4, 2006

Outline

Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

We examine optimal control problems for

convection-diffusion-reaction equations

$$\frac{\partial}{\partial t}\mathbf{x} + \nabla \cdot (\mathbf{C}(\mathbf{x}) - \mathbf{K}(\nabla \mathbf{x})) - \mathbf{Q}(\mathbf{x}) = \mathcal{B}\mathbf{u}(t), \qquad t \in [0, T_f], \quad (1)$$

on $\Omega \in \mathbb{R}^d$, $d = 1, 2, 3$, with appropriate initial and boundary data. Here $\mathbf{C}$ is the convective part, $\mathbf{K}$ is the diffusive part and $\mathbf{Q}$ is the reactive part.

After variational formulation and linearization these can be written as

abstract Cauchy problems

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t). \quad (2)$$

The state space $\mathcal{X}$ and the control space $\mathcal{U}$ are considered Hilbert spaces (e.g. $\mathcal{X} = H^2(\Omega)$ and $\mathcal{U} = L^2(\Omega)$ for the Dirichlet problem for the heat equation with distributed/point control on the unit square).

Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

Only certain measurements of $\mathbf{x}$ available as outputs $\mathbf{y} \in \mathcal{Y} \Rightarrow$

output equation

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t). \tag{3}$$

The linear-quadratic regulator (LQR) problem for (1) can then be expressed as

LQR problem for the abstract Cauchy equation

Minimize the quadratic cost function

$$J(\mathbf{u}) = \int_0^\infty < \mathbf{y}, \mathbf{Q}\mathbf{y} >_{\mathcal{Y}} + < \mathbf{u}, \mathbf{R}\mathbf{u} >_{\mathcal{U}} \, dt, \tag{4}$$

with respect to the linear constraints (2), (3).

Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

[Gibson '78, Balakrishnan '77, Lasiecka/Triggiani '00] discuss that under suitable conditions on $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{Q}$ and $\mathbf{R}$, $\mathbf{u}$ is given as the

optimal feedback

$$\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^*\mathbf{P}_\infty \mathbf{x}(t), \tag{5}$$

where $\mathbf{P}_\infty = \mathbf{P}_\infty^*$ is the minimal solution of the

algebraic operator Riccati equation

$$0 = \mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^*\mathbf{P} + \mathbf{C}\mathbf{Q}\mathbf{C} =: \mathfrak{R}(\mathbf{P}). \tag{6}$$

Note: $0 = \mathfrak{R}(\mathbf{P}) \;\Leftrightarrow\; < \mathbf{v}, \mathfrak{R}(\mathbf{P})\mathbf{w} >= 0 \quad \forall \mathbf{v}, \mathbf{w} \in \operatorname{dom} \mathbf{A}$

Problem background in control of parabolic PDEs

Semidiscretization and finite dimensional systems

Spatial semidiscretization of (1): $\mathcal{X} \rightsquigarrow \mathcal{X}_h$ in (2) and $\mathcal{Y} \rightsquigarrow \mathcal{Y}_h$ in (3) $\Rightarrow$ large scale sparse ODE system

$$\dot{x}_h = A_h x_h + B_h u, \qquad y_h = C_h x_h$$

with cost function

$$J_h(u) = \int_0^\infty < y_h, Q_h y_h > + < u, Ru > dt,$$

and u is given in feedback form as

$$u = -R^{-1} B_h^T P_h x_h$$

where P_h is the minimal selfadjoint solution of the algebraic matrix Riccati equation (ARE)

$$0 = A_h^T P + P A_h - P B_h R^{-1} B_h^T P + C_h Q_h C_h =: \mathfrak{R}_h(P). \tag{7}$$

Problem background in control of parabolic PDEs

Semidiscretization and finite dimensional systems

Spatial semidiscretization of (1): $\mathcal{X} \rightsquigarrow \mathcal{X}_h$ in (2) and $\mathcal{Y} \rightsquigarrow \mathcal{Y}_h$ in (3) $\Rightarrow$ large scale sparse ODE system

$$M\dot{x}_h = A_h x_h + B_h u, \qquad y_h = C_h x_h$$

with cost function

$$J_h(u) = \int_0^\infty < y_h, Q_h y_h > + < u, Ru > dt,$$

and u is given in feedback form as

$$u = -R^{-1} B_h^T P_h M x_h$$

where P_h is the minimal selfadjoint solution of the algebraic matrix Riccati equation (ARE)

$$0 = A_h^T P M + M P A_h - M P B_h R^{-1} B_h^T P M + C_h Q_h C_h =: \mathfrak{R}_h(P). \qquad (7)$$

Problem background in control of parabolic PDEs

Approximation of the ∞-dim. system

Under some natural conditions on the discretization (e.g. Galerkin scheme) it can be shown

Theorem [Banks/Kunisch'84, Lasiecka/Triggiani '00, Curtain '03]

The n-dim. problems approximate the abstract Cauchy problem:

- $P_h \Pi_h \mathbf{v} \rightarrow P\mathbf{v}$ as $h \rightarrow 0$ for all $\mathbf{v} \in \mathcal{X}$,
- $S_h(t) \Pi_h \mathbf{v} \rightarrow S(t)\mathbf{v}$ as $h \rightarrow 0$ for all $\mathbf{v} \in \mathcal{X}$,

where $S_h(t)$ and $S(t)$ are the solution semigroups generated by $(A_h - B_h R^{-1} B_h^T P_h)$, $(A - BR^{-1}B^T P)$ respectively.

Remark

Note $\mathbf{u}$ for n-dim. and ∞-dim. problems from the same function space.

- no discretization of the controls required
- simulated controls can directly be applied to the ∞-dim. system

LRCF Newton Method for the ARE

Newton's method for solving the ARE

Newton's iteration for the ARE

$$\mathfrak{R}_h'|_P(N_l) = -\mathfrak{R}_h(P_l), \qquad P_{l+1} = P_l + N_l,$$

where the Frechét derivative of $\mathfrak{R}_h$ at P is the Lyapunov operator

$$\mathfrak{R}_h'|_P : \quad Q \mapsto (A_h - B_h R^{-1} B_h^T P)^T Q + Q(A_h - B_h R^{-1} B_h^T P),$$

can be rewritten as

one iteration step

$$(A_h - B_h R^{-1} B_h^T P_l)^T P_{l+1} + P_{l+1}(A_h - B_h R^{-1} B_h^T P_l) = -C_h^T Q_h C_h - P_l B_h R^{-1} B_h^T P_l$$

i.e. in every Newton step we have to solve a

Lyapunov equation

$$F^T P + PF = -GG^T. \tag{8}$$

LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

Peaceman Rachford ADI:

Consider $Au = s$ where $A \in \mathbb{R}^{n\times n}$ spd, $s \in \mathbb{R}^n$. ADI Iteration Idea:
Decompose $A = H + V$ with $H, V \in \mathbb{R}^{n\times n}$ such that

$$(H + pI)v = r$$
$$(V + pI)w = t$$

can be solved easily.

ADI Iteration

If H, V spd $\Rightarrow \exists p_j,\ j = 1, 2, \ldots J$ such that

$$\begin{aligned} u_0 &= 0 \\ (H + p_j I)u_{j-\frac{1}{2}} &= (p_j I - V)u_{j-1} + s \\ (V + p_j I)u_j &= (p_j I - H)u_{j-\frac{1}{2}} + s \end{aligned} \qquad (9)$$

converges to $u \in \mathbb{R}^n$ solving $Au = s$.

LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

Peaceman Rachford ADI:

Consider $Au = s$ where $A \in \mathbb{R}^{n \times n}$ spd, $s \in \mathbb{R}^n$. ADI Iteration Idea:
Decompose $A = H + V$ with $H, V \in \mathbb{R}^{n \times n}$ such that

$$\begin{aligned}(H + pI)v &= r \\ (V + pI)w &= t\end{aligned}$$

can be solved easily.

ADI Iteration

If H, V spd $\Rightarrow \exists p_j$, $j = 1, 2, \ldots J$ such that

$$\begin{aligned} u_0 &= 0 \\ (H + p_j I)u_{j-\frac{1}{2}} &= (p_j I - V)u_{j-1} + s \\ (V + p_j I)u_j &= (p_j I - H)u_{j-\frac{1}{2}} + s \end{aligned} \tag{9}$$

converges to $u \in \mathbb{R}^n$ solving $Au = s$.

LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

The Lyapunov operator

$$\mathcal{L}: \quad P \quad \mapsto \quad F^T P + PF$$

can be decomposed into the linear operators

$$\mathcal{L}_H : P \mapsto F^T P \qquad \mathcal{L}_V : P \mapsto PF.$$

Such that in analogy to (4) we find the

ADI iteration for the Lyapunov equation (8)

$$\begin{aligned} P_0 &= 0 \\ (F^T + p_j I) P_{j-\frac{1}{2}} &= -GG^T - P_{j-1}(F - p_j I) \\ (F^T + p_j I) P_j^T &= -GG^T - P_{j-\frac{1}{2}}^T (F - p_j I) \end{aligned}$$

- Can be rewritten to iterate on the low rank Cholesky factors Z_j of P_j to exploit $\mathrm{rk}(P_j) \ll n$. [Li/White 2002; Penzl 1999; Benner/Li/Penzl 2000]

ADI shift parameter computation

The ADI Min-Max-problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j \in \mathbb{R} | j=1,\dots,J\} \subset \mathbb{R}} \quad \max_{\lambda \in \sigma(H), \gamma \in \sigma(V)} \quad \left| \prod_{j=1}^{J} \frac{(p_j - \lambda)(p_j - \gamma)}{(p_j + \lambda)(p_j + \gamma)} \right| .$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different computational methods for the complex case around 1990.

ADI shift parameter computation

The ADI Min-Max-problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j \in \mathbb{R} | j=1,\dots,J\} \subset \mathbb{R}} \quad \max_{\lambda \in \sigma(F),} \quad \left| \prod_{j=1}^{J} \frac{(p_j - \lambda)}{(p_j + \lambda)} \right| .$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different computational methods for the complex case around 1990.

ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

1. optimal parameters [WACHSPRESS: ADI model problem '95] :
   - Solve the min-max-problem on an elliptic functions region.
   - Computation needs knowledge of the complete spectrum of F.
2. heuristic parameters [PENZL: Lyapack '99]
   - use approximated eigenvalues as shifts
   - suboptimal $\Rightarrow$ convergence might be weak
3. semi-optimal parameters: Idea: combine the advantages of these methods
   - use Arnoldi's method to approximate the outer spectrum
   - compute the optimal parameters for this approximation

ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

1. optimal parameters [WACHSPRESS: ADI model problem '95] :
   - Solve the min-max-problem on an elliptic functions region.
   - Computation needs knowledge of the complete spectrum of F.
2. heuristic parameters [PENZL: Lyapack '99]
   - use approximated eigenvalues as shifts
   - suboptimal $\Rightarrow$ convergence might be weak
3. semi-optimal parameters: Idea: combine the advantages of these methods
   - use Arnoldi's method to approximate the outer spectrum
   - compute the optimal parameters for this approximation

ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

1. optimal parameters [WACHSPRESS: ADI model problem '95] :
    - Solve the min-max-problem on an elliptic functions region.
    - Computation needs knowledge of the complete spectrum of F.
2. heuristic parameters [PENZL: Lyapack '99]
    - use approximated eigenvalues as shifts
    - suboptimal $\Rightarrow$ convergence might be weak
3. semi-optimal parameters: Idea: combine the advantages of these methods
    - use Arnoldi's method to approximate the outer spectrum
    - compute the optimal parameters for this approximation

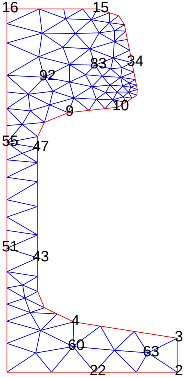
Numerical results

The model problem

- Mathematical model: boundary control for linearized 2D heat equation.

$$
\begin{aligned}
c \cdot \rho \frac{\partial}{\partial t} x &= \lambda \Delta x, && \xi \in \Omega \\
\lambda \frac{\partial}{\partial n} x &= \kappa(u_k - x), && \xi \in \Gamma_k,\ 1 \leq k \leq 7, \\
\frac{\partial}{\partial n} x &= 0, && \xi \in \Gamma_7.
\end{aligned}
$$

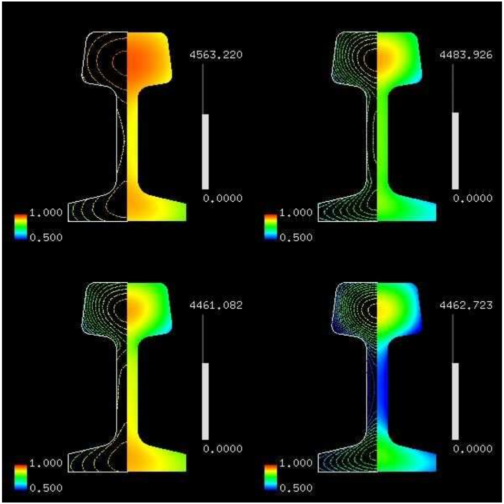
- FEM discretization, different models for initial mesh ($n = 371$), 1, 2, 3, 4 steps of mesh refinement $\Rightarrow$ $n = 1357, 5177, 20209, 79841$.



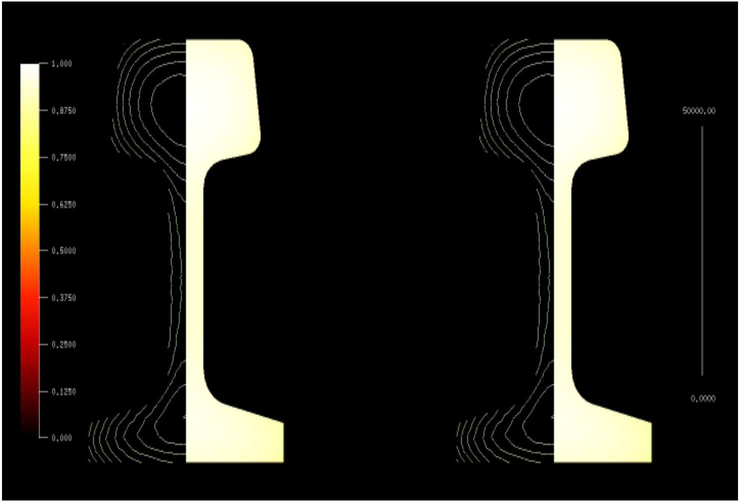
Source: Physical model: courtesy of Mannesmann/Demag.

Math. model: Tröltzsch/Unger 1999/2001, Penzl 1999, S. 2003.

Numerical results
Stabilization



Numerical results
Tracking



Numerical results

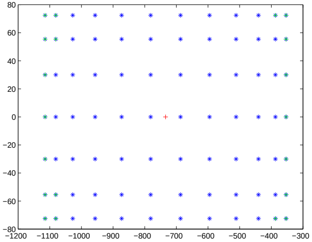
Parameter comparison

$$x_t = \Delta x + \begin{pmatrix} 20 \\ 0 \end{pmatrix} . \nabla x - 180x + Bu$$

on $\Omega = \{(0,1) \times (0,1)\} \times (0, \infty)$. Finite difference semidiscretization leads to a system of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned}$$

where A is nonsymmetric and its spectrum is complex.



Source: Slicot Working Note 2002-2: A collection of Benchmark examples for model reduction of linear time invariant dynamical systems [Chahlaoui/Van Dooren 2002]

Numerical results

Parameter comparison

Sizes of the low rank Cholesky factors:

heuristic: m=262

optimal/ semi-optimal: m=76

all of them are real in this case.