Numerical Solution of Linear-Quadratic Optimal Control Problems for Parabolic PDEs

Peter Benner

Zentrum für Technomathematik
Fachbereich 3 – Mathematik und Informatik
Universität Bremen

SIAM–EMS Conference
Berlin, September 2–6, 2001

Dedicated to Thilo Penzl

Outline

- Linear-quadratic optimal control of parabolic PDEs
- Algebraic Riccati equations and their numerical solution
- A low-rank Newton method
- Numerical examples
- Conclusions and outlook

LQ Optimal Control of Parabolic PDEs

Linear parabolic PDE (e.g., heat equation, convection-diffusion equation):

$$\frac{\partial x}{\partial t} - \nabla\left(A(\xi)\nabla x\right) + d(\xi)\nabla x + r(\xi)x = Bu(t),$$
$$\xi \in \Omega,\ t > 0,$$

with initial and boundary conditions ($\partial\Omega = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$)

$$\begin{aligned} x(\xi,0) &= x_0(\xi), \quad \xi \in \Omega, \\ x(\xi,t) &= B_1 u_1(t), \qquad \xi \in \Gamma_1, \\ \frac{\partial}{\partial \eta} x(\xi,t) &= B_2 u_2(t), \qquad \xi \in \Gamma_2, \\ x(\xi,t) + \frac{\partial}{\partial \eta} x(\xi,t) &= B_3 u_3(t), \qquad \xi \in \Gamma_3. \end{aligned}$$

- $B = 0 \implies$ boundary control problem
- $B_j = 0\ \forall j \implies$ point control problem

Output equation:

$$y = Cx, \qquad t \geq 0.$$

Quadratic performance index:

$$\min_u \mathcal{J}(x_0, u) = \frac{1}{2}\int_0^\infty \left(\|y\|_{\mathcal{Y}}^2 + \|u\|_{\mathcal{U}}^2\right) dt,$$

Abstract Setting: Linear-Quadratic Regulator Problem

Given Hilbert spaces

- $\mathbb{X}$ – state space,
- $\mathbb{U}$ – control space,
- $\mathbb{Y}$ – output space,

and operators

$$\mathcal{A} : \mathrm{dom}(\mathcal{A}) \subset \mathbb{X} \to \mathbb{X}, \quad \mathcal{B} : \mathbb{U} \to \mathbb{X}, \quad \mathcal{C} : \mathbb{X} \to \mathbb{Y}.$$

LQR Problem:

Minimize

$$\mathcal{J}(x_0, u) = \frac{1}{2} \int_0^\infty \left(\|y\|_{\mathcal{Y}}^2 + \|u\|_{\mathcal{U}}^2 \right) dt,$$

for $u \in \mathbb{L}_2(0, \infty; \mathbb{U})$, where

$$\dot{x} = \mathcal{A}x + \mathcal{B}u, \qquad x(0) = x_0 \in \mathbb{X},$$

$$y = \mathcal{C}x.$$

Example

Heat equation with point control:

$$\begin{aligned} x_t &= \Delta x + b(\xi)u(t) \text{ in } \Omega, \qquad x = 0 \text{ on } \delta\Omega, \\ y &= \int_\Omega c(\xi)x \, d\xi \end{aligned}$$

Weak formulation with test functions $v \in \mathbb{H}_0^1(\Omega)$:

$$\begin{aligned} \int_\Omega x_t v \, d\xi &= \int_\Omega \Delta x v \, d\xi + \int_\Omega b(\xi)u(t)v \, d\xi \\ &= -\int_\Omega \nabla x \nabla v \, d\xi + \left(\int_\Omega bv \, d\xi \right) u(t) \end{aligned}$$

Then $\mathbb{X} = \mathbb{L}_2(\Omega)$, $\mathbb{U} = \mathbb{R} = \mathbb{Y}$, and with

$$< w, v > := \int_\Omega wv \, d\xi$$

define linear operators:

$$\begin{aligned} < \mathcal{A}w, v > &:= -\int_\Omega \nabla w \nabla v \, d\xi \\ \mathcal{B}u &:= b(\xi)u(t) \\ \mathcal{C}v &:= \int_\Omega c(\xi)v \, d\xi \end{aligned}$$

Solution of the LQR Problem

Theorem [*Gibson '79*]

Assumptions:

- $\mathcal{A}$ infinitesimal generator of C_0-semigroup.
- $\mathcal{B}, \mathcal{C}$ linear, bounded.
- $(\mathcal{A}, \mathcal{B})$ stabilizable ($\exists \mathcal{K} : \mathbb{X} \to \mathbb{U}$ linear, bounded, such that C_0-semigroup generated by $\mathcal{A} - \mathcal{B}\mathcal{K}$ is exponentially stable.)
- $(\mathcal{C}, \mathcal{A})$ detectable, i.e., $(\mathcal{A}^*, \mathcal{C}^*)$ stabilizable.
- $\forall x_0 \in \mathbb{X}$ there exists admissible control u.
  ($u \in \mathbb{L}_2(0, \infty; \mathbb{U})$ admissible $\iff$ $\mathcal{J}(x_0, u) < \infty$.)

Then: The algebraic operator Riccati equation

$$0 = \mathcal{R}(\mathcal{P}) := \mathcal{C}^*\mathcal{C} + \mathcal{A}^*\mathcal{P} + \mathcal{P}\mathcal{A} - \mathcal{P}\mathcal{B}\mathcal{B}^*\mathcal{P}$$

has unique, selfadjoint solution $\mathcal{P}_\infty$, where

- $\mathcal{P}_\infty : \text{dom}(\mathcal{A}) \to \text{dom}(\mathcal{A}^*)$ linear, bounded,
- $\mathcal{P}_\infty \geq 0$, i.e., positive semidefinite.

Solution of LQR problem is feedback control:

$$u_\infty(t) = -\mathcal{B}^*\mathcal{P}_\infty x(t) = -\mathcal{K}_\infty x(t).$$

$\mathcal{P}_\infty$ is stabilizing, that is, the C_0-semigroup generated by $\mathcal{A} - \mathcal{B}\mathcal{B}^*\mathcal{P}_\infty$ is exponentially stable.

Numerical Solution

Galerkin approach, space discretization by finite element method $\Rightarrow$ solve LQR problem on $\mathbb{X}_n \subset \mathbb{X}$, $\dim(\mathbb{X}_n) = n$:

Minimize

$$\mathcal{J}(x_0, u) = \frac{1}{2}\int_0^\infty \left(y^T y + u^T u\right) dt,$$

for $u \in \mathbb{L}_2(0, \infty; \mathbb{R}^m)$, where

$$\begin{aligned} M\dot{x} &= -Lx + Bu, \qquad x(0) = x_0, \\ y &= Cx, \end{aligned}$$

with stiffness matrix $L \in \mathbb{R}^{n\times n}$, mass matrix $M \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$.

Solution of finite-dimensional LQR problem given by

$$u_*(t) = -B^T P_* x(t) =: -K_* x(t),$$

where $P_* \geq 0$ is stabilizing solution of the algebraic Riccati equation (ARE)

$$0 = \mathcal{R}(P) := C^T C + A^T P + PA - PBB^T P,$$

with $A := -M^{-1}L$, $B := M^{-1}B$.

Convergence: *Gibson '79, Banks/Kunisch '84, Lasiecka/Triggiani '91*

Algebraic Riccati Equations

General form:

$$0 = \mathcal{R}(P) := Q + A^T P + PA - PGP$$

with given $A, G, Q \in \mathbb{R}^{n\times n}$ and unknown $P \in \mathbb{R}^{n\times n}$.

Symmetric ARE: $G = G^T$, $Q = Q^T$.

Here, control-theoretic assumptions ensure existence of unique stabilizing solution P_*, i.e.,

$$\sigma\left(A - GP_*\right) \subset \mathbb{C}^-.$$

(In LQR problems, $P_* = P_*^T \geq 0$.)

In large scale applications from semi-discretized control problems for PDEs,

- $n = 10^3 - 10^5$ ($\Longrightarrow$ $10^6 - 10^{10}$ unknowns!),
- A has sparse representation,
- G, Q low-rank with
  - $G = BB^T$, $B \in \mathbb{R}^{n\times m}$, $m \ll n$,
  - $Q = C^T C$, $C \in \mathbb{R}^{p\times n}$, $p \ll n$.

Numerical Solution of AREs

First approach: *[Potter '66, Laub '79,...]*

Use connection to Hamiltonian eigenproblem.

P is stabilizing solution of the ARE

$$\Longleftrightarrow$$

$$\begin{aligned} H \begin{bmatrix} I_n \\ P \end{bmatrix} = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} I_n \\ P \end{bmatrix} &= \begin{bmatrix} I_n \\ P \end{bmatrix} (A - GP), \\ \sigma(A - GP) &= \sigma(H) \cap \mathbb{C}^- \end{aligned}$$

I.e., columns of $\begin{bmatrix} I_n \\ P \end{bmatrix}$ span stable invariant subspace of Hamiltonian Matrix H.

Note: here, $\sigma(H) = \{\pm\lambda_j \,|\mathrm{Re}(\lambda_j) < 0\}$.

Definition:

$H \in \mathbb{R}^{2n\times 2n}$ Hamiltonian

$$\Longleftrightarrow$$

$HJ = (HJ)^T$, where $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$, in other words, H is skew-symmetric w.r.t. $\langle x, y\rangle_J = x^T J y$.

Methods:

- Compute stable H-invariant subspace via (structured, block-) Schur decomposition,

$$T^{-1}HT = \begin{bmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{bmatrix}, \quad \sigma(H_{11}) = \sigma(H) \cap \mathbb{C}^-,$$

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \Longrightarrow P = T_{21}T_{11}^{-1}$$

  - QR algorithm [*Laub '79*];
  - SR algorithm [*Bunse-Gerstner/Mehrmann '86*];
  - multishift algorithm [*Ammar/B./Mehrmann '93*];
  - embedding algorithm [*B./Mehrmann/Xu '97*];

  or spectral projection methods,

  - sign function method [*Roberts '71, Byers '87, Gardiner/Laub '86*]
  - disk function method [*Malyshev '93, Bai/Demmel/Gu '95, B./Byers '95, B. '97*]

  $\Longrightarrow$ $\mathcal{O}(n^3)$, sparse matrix structure is destroyed.

- Krylov subspace methods $\Rightarrow$ employ sparse matrix structure, but need n-dimensional subspace!

Newton's Method for AREs

Other approach:

Consider

$$0 = \mathcal{R}(P) = C^T C + A^T P + PA - PBB^T P$$

with stable A, i.e., $\sigma(A) \subset \mathbb{C}^-$, as nonlinear system of equations.

Frechét derivative of $\mathcal{R}(P)$ at P:

$$\mathcal{R}'_P : Z \to (A - BB^T P)^T Z + Z(A - BB^T P)$$

Newton-Kantorovich method:

$$P_{j+1} = P_j - \left(\mathcal{R}'_{P_j}\right)^{-1} \mathcal{R}(P_j), \qquad j = 0,\ 1,\ 2,\ \ldots$$

$\Longrightarrow$ Newton's method (with line search) for AREs
[*Kleinman '68, Mehrmann '91, Lancaster/Rodman '95, B./Byers '94/'98, B. '97, Guo/Laub '99*]

1. $P_0 = 0.$

2. FOR $j = 0, 1, 2, \ldots$

   2.1 $A_j \leftarrow A - BB^T P_j =: A - BK_j.$

   2.2 Solve Lyapunov equation

   $$A_j^T N_j + N_j A_j = -\mathcal{R}(P_j).$$

   2.3 $P_{j+1} \leftarrow P_j + t_j N_j.$

   END FOR j

Properties

- Choice of t_j via solution of minimization problem corresponding to $\mathcal{R}(P) = 0$ (exact line search):

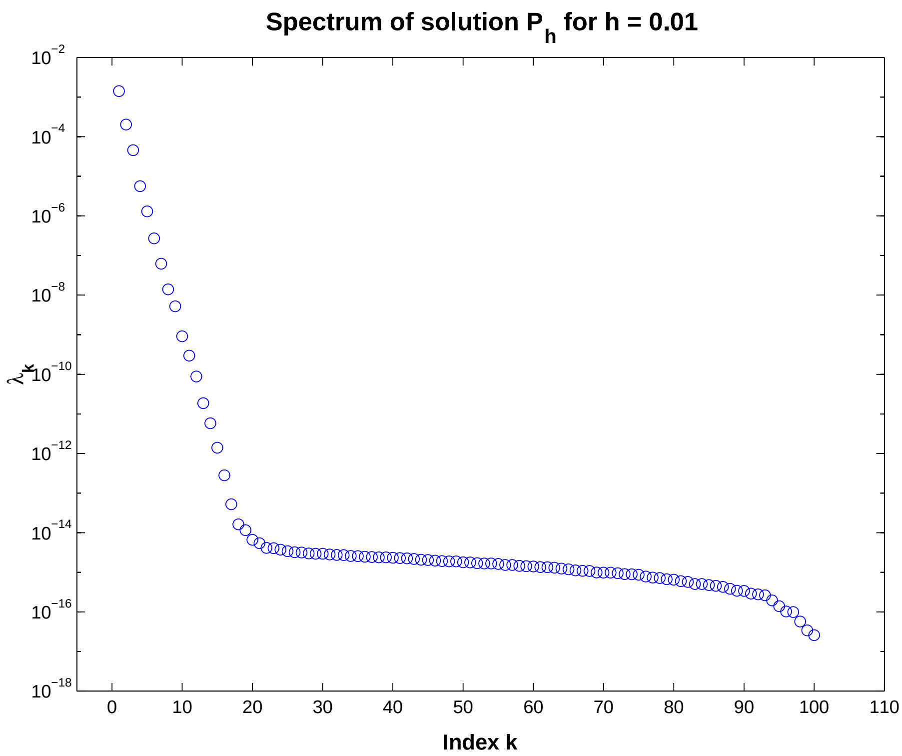
$$\begin{aligned} \min_t f(t) &= \min_t \|\mathcal{R}(P + tN)\|_F^2 \\ &= \min_t \operatorname{trace} \left(\mathcal{R}(P + tN)^T \mathcal{R}(P + tN) \right). \end{aligned}$$

- Convergence:
  - $A_j = A - BK_j = A - BB^T P_j$ is stable $\forall\, j \geq 1$.
  - $\|\mathcal{R}(P_j)\|_F \geq \|\mathcal{R}(P_{j+1})\|_F \;\forall\, j \geq 0$.
  - $\lim_{j\to\infty} \|\mathcal{R}(P_j)\|_F = 0$.
  - $P_* \leq \ldots \leq P_{j+1} \leq P_j \leq \ldots \leq P_1$ (if $t_j \equiv 1$).
  - $\lim_{j\to\infty} P_j = P_* \geq 0$ (locally quadratic).
- Need sparse Lyapunov solver.
- BUT: $P = P^T \in \mathbb{R}^{n\times n} \Longrightarrow n(n+1)/2$ unknowns!

Low-Rank Approximation

Consider spectrum of ARE solution.

Example: Linear 1D heat equation with point control,
$\Omega = [\,0,\ 1\,]$, FEM discretization using linear B-splines,
$h = 1/100$ ($\Longrightarrow n = 101$).



Idea:

$$P = P^T \geq 0 \quad \Longrightarrow \quad P = ZZ^T = \sum_{k=1}^{n} \lambda_k z_k z_k^T$$

$$\lambda_k \approx 0,\ k > r \quad \Longrightarrow \quad P \approx Z^{(r)}(Z^{(r)})^T = \sum_{k=1}^{r} \lambda_k z_k z_k^T.$$

Iteration for $Z^{(r)}$

Re-write Newton's method for AREs [*Kleinman '68*]

$$A_j^T N_j + N_j A_j = -\mathcal{R}(P_j)$$

$$\Longleftrightarrow$$

$$A_j^T \underbrace{(P_j + N_j)}_{=P_{j+1}} + \underbrace{(P_j + N_j)}_{=P_{j+1}} A_j = \underbrace{-C^T C - P_j B B^T P_j}_{=: -W_j W_j^T}$$

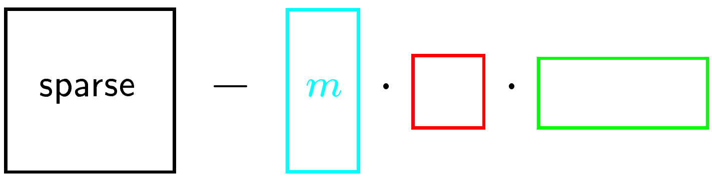
Set $P_j = Z_j Z_j^T$ for $\operatorname{rank}(Z_j) \ll n$:

$$A_j^T \left(Z_{j+1} Z_{j+1}^T\right) + \left(Z_{j+1} Z_{j+1}^T\right) A_j = -W_j W_j^T$$

$$\Downarrow$$

Solve Lyapunov equations for Z_{j+1} directly and use *'sparse + low-rank'* structure of A_j,

$$A_j = A - BK_j \;=\; A \;-\; B \cdot (B^T Z_j) \cdot Z_j^T,$$

$=$ 

$m \ll n \Longrightarrow$ "inversion" using Sherman-Morrison-Woodbury formula:

$$(A - BK_j)^{-1} = (I_n + A^{-1}B(I_m - K_j A^{-1} B)^{-1} K_j) A^{-1}.$$

ADI-Method for Lyapunov equations

[*Wachspress '88*]

Let $A \in \mathbb{R}^{n \times n}$ be stable ($\sigma(A) \in \mathbb{C}^-$), $W \in \mathbb{R}^{n \times w}$ ($w \ll n$), consider Lyapunov equation

$$A^T Q + QA = -WW^T.$$

ADI iteration:

$$\begin{aligned} (A^T + p_k I) Q_{(k-1)/2} &= -WW^T - Q_{k-1}(A - p_k I) \\ (A^T + \overline{p_k} I) {Q_k}^T &= -WW^T - Q_{(k-1)/2}(A - \overline{p_k} I) \end{aligned}$$

with parameters $p_k \in \mathbb{C}^-$ and $p_{k+1} = \overline{p_k}$ in case $p_k \notin \mathbb{R}$.

With $Q_0 = 0$ and appropriate choice of p_k:

$$\lim_{k \to \infty} Q_k = Q \text{ superlinear.}$$

Factored ADI Iteration

[*B./Li/Penzl '00*]

Set $Q_k = Y_k Y_k^T$, re-formulation $\Longrightarrow$

$$V_1 \quad \leftarrow \quad \sqrt{-2\mathrm{Re}\,(p_1)}(A + p_1 I)^{-1} W$$

$$Y_1 \quad \leftarrow \quad V_1$$

FOR $k = 2, 3, \ldots$

$$V_k \leftarrow \frac{\sqrt{\mathrm{Re}\,(p_k)}}{\sqrt{\mathrm{Re}\,(p_{k-1})}} \left(I - (p_k + \overline{p_{k-1}})(A + p_k I)^{-1} \right) V_{k-1}$$

$$Y_k \leftarrow \left[\begin{array}{cc} Y_{k-1} & V_k \end{array} \right]$$

$\Longrightarrow$

$$Y_{k_{\max}} = \left[\begin{array}{ccc} V_1 & \ldots & V_{k_{\max}} \end{array} \right],$$

with

$$V_k \;=\; \big[\;\big] \in \mathbb{C}^{n \times w}$$

and

$$Y_{k_{\max}} Y_{k_{\max}}^T \approx Q.$$

Newton-ADI for AREs

[*B./Li/Penzl '00*]

Solve Lyapunov equation

$$(A-BK_{j-1})^T Z_j Z_j^T + Z_j Z_j^T (A-BK_{j-1}) = -W_{j-1}W_{j-1}^T$$

with factored ADI iteration

$$\Downarrow$$

Sequence $Y_0^{(j)}, Y_1^{(j)}, \ldots, Y_{k_{\max}}^{(j)}$ of low-rank approximations to solution of Lyapunov equation

$$\Downarrow$$

$$Z_j = Y_{k_{\max}}^{(j)}$$

$$\Downarrow$$

Newton's method with factored iterates

$$P_j = Z_j Z_j^T$$

$$\Downarrow$$

Factored solution of ARE:

$$P_* \approx Z_{j_{\max}} Z_{j_{\max}}^T$$

Solution of LQR Problems

Recall: solve LQR problem via ARE.

But: ARE is detour, need feedback!

$$K = B^T P = B^T Z Z^T$$

Idea: Direct iteration for feedback matrix.

- Approximate feedback matrix in step j of Newton iteration:

$$K_j = B^T Z_j Z_j^T = \sum_{k=1}^{k_{\max}} (B^T V_{j,k}) V_{j,k}^T$$

- Direct updating inside ADI iteration possible:

$$K_{j,0} = 0, \quad K_{j,k} = K_{j,k-1} + (B^T V_{j,k}) V_{j,k}^T$$

- Set $K := K_{j_{\max},k_{\max}}$.

- Requires only workspace of size $m \times n$ for feedback matrix and $n \times (m + p)$ for $V_{j,k}$.

Numerical Examples

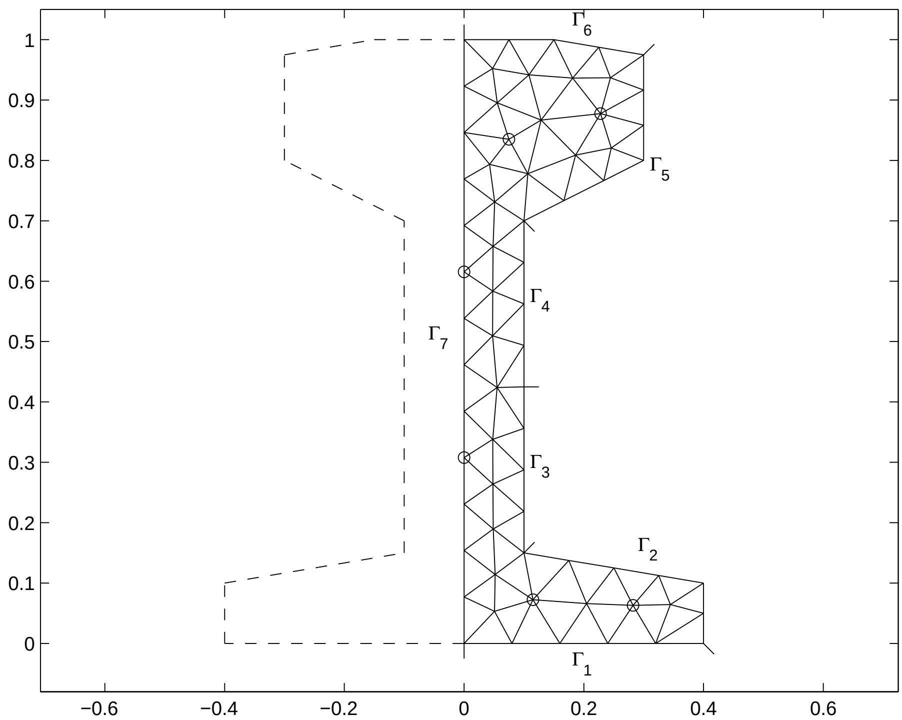
Example 1 [*Tröltzsch/Unger '99, Penzl '99*]

- Optimal cooling of steel profiles.
- Model: boundary control for linearization of 2D heat equation.

$$\begin{aligned} x_t &= \Delta x, && x \in \Omega \\ x + x_\eta &= u_k, && x \in \Gamma_k, \; k = 1, \ldots, 6, \\ x_\eta &= 0, && x \in \Gamma_7. \end{aligned}$$

$\Longrightarrow m = p = 6$

- FEM discretization, initial mesh ($n = 821$).



2 refinement steps $\Longrightarrow n = 3113$.

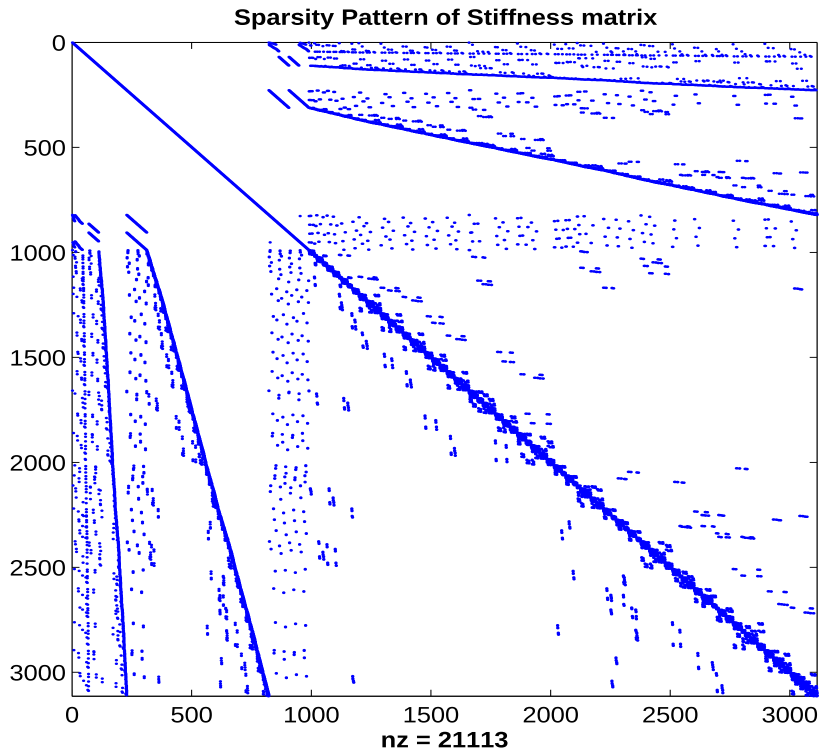
Numerical Examples

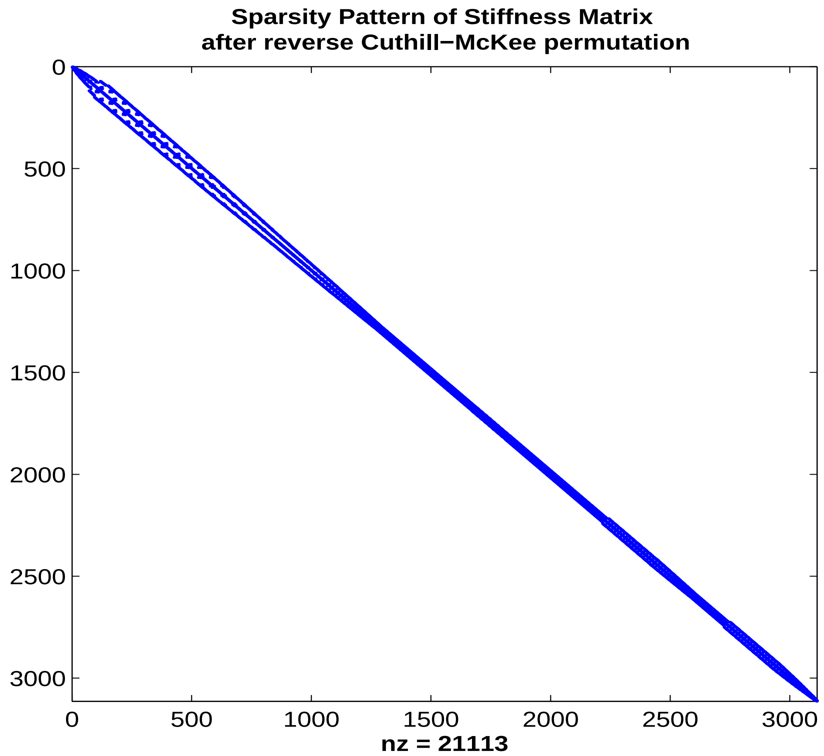
Solution of linear systems of equations:

- Instead of $A = -M^{-1}L$ consider $A = -M_C^{-1}LM_C^{-T}$, M_C = Cholesky factor of M,
- Cholesky factorization and solution of 'shifted' linear systems using sparse direct solver.

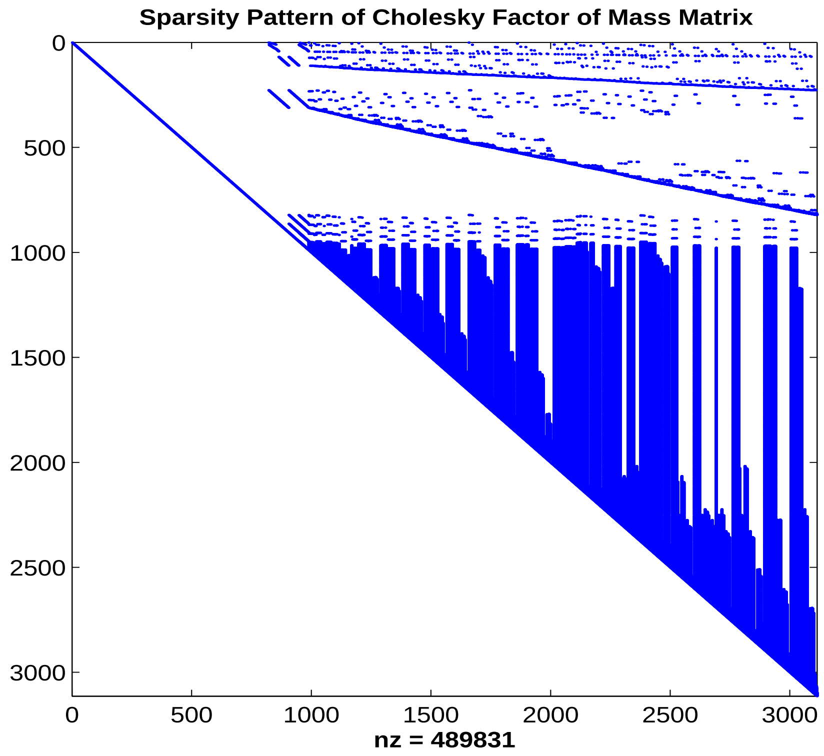
Example		1a	1b	1c
ARE	Newton iterations	5	8	12
	# columns of $\tilde{Z}$	540	492	522
	$\frac{\|\mathcal{R}(\tilde{Z}\tilde{Z}^H)\|_F}{\|C^TC\|_F}$	$7 \cdot 10^{-14}$	$4 \cdot 10^{-14}$	$1 \cdot 10^{-14}$
Lyapunov eq.	min. # iterations	45	40	42
	max. # iterations	46	45	46

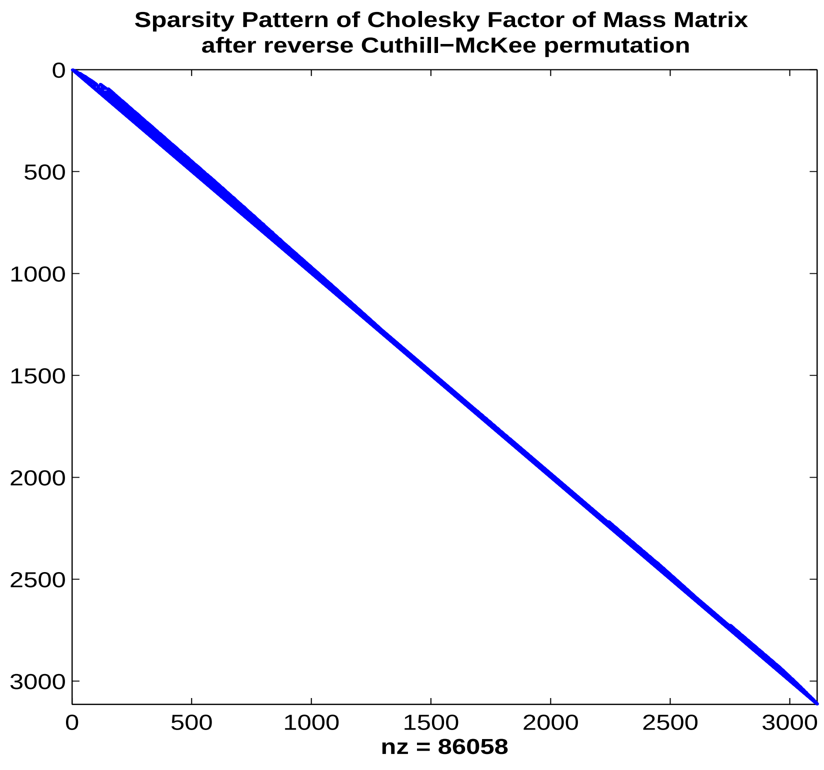
Example 1, Stiffness Matrix





Example 1, Cholesky Factor of Mass Matrix





Example 2: Direct Feedback Iteration

Test scalability:

- Linear 3D convection-diffusion equation with boundary control in unit cube.
- Finite differences discretization on uniform grid.
- Solution of linear systems of equations using QMR and ILU preconditioning.

Example		2a	2b	2c
(n, m, p)		(1000,1,1)	(5832,1,1)	(27000,1,1)
feedback	Newton iterations	4	4	3
	$\frac{\|\tilde{K}-K\|_F}{\|K\|_F}$	$1.3 \cdot 10^{-8}$	$8.8 \cdot 10^{-8}$	–
Lyapunov eq.	min. # iterations	103	143	96
	max. # iterations	129	143	96

Conclusions and Outlook

- Solution of LQR problems for parabolic PDEs via low-rank factor ADI-Newton method is efficient and reliable.
- Riccati-approach applicable to other control problems for linear evolution equations as well.
- Newton's method guarantees stabilization property of low-rank ARE solutions!
- Direct computation of feedback matrix for LQR problem possible without ARE detour.
- Number of columns in low-rank factors can be kept low using column compression with updating technique.
- Need analysis on how accurate Lyapunov equations need to be solved (inexact Newton methods).
- Line search for ADI-Newton method efficient (i.e. reduces no. of iterations), but too expensive (w.r.t. flops per step).