Rosenbrock methods for solving differential Riccati equations

Hermann Mena
joint work with
Peter Benner

Mathematik in Industrie und Technik
Fakltät für Mathematik
TU-Chemnitz
Departamento de Matemática
Escuela Politécnica Nacional
Quito, Ecuador

June 15, 2007

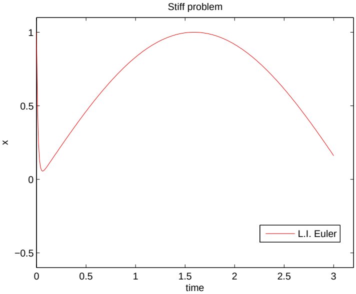
Outline

Rosenbrock Methods

Introduction

Features:

- Excellent stability properties.
- Easy to implement.
- Suitable for parallelization.
- Applicable to $M\dot{x} = f(x)$.
- Rapid change of step size.
- Have already proven be very effective (ODEs), e.g., chemical kinetics.



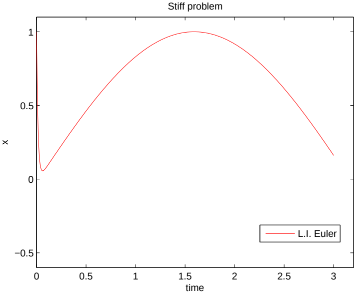
$$\dot{x}(t) = -100(x(t) - sin(t)), \qquad x(0) = 1 \quad t \geq 0.$$

Rosenbrock Methods

Introduction

Features:

- Excellent stability properties.
- Easy to implement.
- Suitable for parallelization.
- Applicable to $M\dot{x} = f(x)$.
- Rapid change of step size.
- Have already proven be very effective (ODEs), e.g., chemical kinetics.



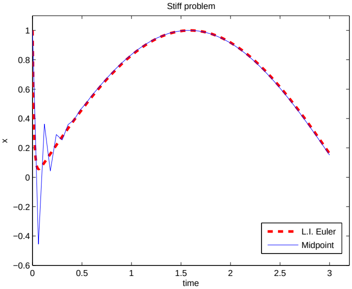
$$\dot{x}(t) = -100(x(t) - sin(t)), \qquad x(0) = 1 \quad t \geq 0.$$

Rosenbrock Methods

Introduction

Features:

- Excellent stability properties.
- Easy to implement.
- Suitable for parallelization.
- Applicable to $M\dot{x} = f(x)$.
- Rapid change of step size.
- Have already proven be very effective, e.g., chemical kinetics.



$$\dot{x}(t) = -100(x(t) - sin(t)), \qquad x(0) = 1 \quad t \geq 0.$$

Rosenbrock methods

Schemes

Let us consider an autonomous ODE system

$$\dot{x}(t) = f(x), \qquad x(t_0) = x_0, \quad t_0 \le t \le t_f.$$

s-stage Rosenbrock method

$$\begin{aligned} k_i &= hf\left(x_n + \textstyle\sum_{j=1}^{i-1} \alpha_{ij} k_j\right) + hJ \textstyle\sum_{j=1}^{i} \gamma_{ij} k_j, \quad i = 1, \dots, s, \\ x_{n+1} &= x_n + \textstyle\sum_{j=1}^{s} b_j k_j, \end{aligned}$$

where α_{ij}, γ_{ij}, b_j are the determining coefficients, h is the step size and J is chosen as $f'(x_n)$.

- For $J = f'(x_0)$ W-methods (T. Steihaug and A. Wolfbrand 1979).
- For $J = 0$ a DIRK method is recovered.

Rosenbrock schemes

Schemes

Let us consider an non-autonomous ODE system

$$\dot{x}(t) = f(x, t), \qquad x(t_0) = x_0, \quad t_0 \le t \le t_f.$$

s-stage Rosenbrock method for non-autonomous systems

$$\begin{aligned} k_i &= hf\left(t_n + \alpha_i h, x_n + \textstyle\sum_{j=1}^{i-1} \alpha_{ij} k_j\right) + hJ \textstyle\sum_{j=1}^{i} \gamma_{ij} k_j \\ &\quad + \gamma_i h^2 \frac{\partial f}{\partial t}(t_n, x_n), \qquad\qquad i = 1, \dots, s, \\ x_{n+1} &= x_n + \textstyle\sum_{j=1}^{s} b_j k_j, \end{aligned}$$

where α_{ij}, γ_{ij}, b_j are the determining coefficients, $J = \frac{\partial f}{\partial t}(x_n)$, and h is the step size, and

$$\alpha_i = \sum_{j=1}^{i-1} \alpha_{ij}, \qquad \gamma_i = \sum_{j=1}^{i} \gamma_{ij}.$$

Rosenbrock schemes

Schemes

Autonomous case

Linearly implicit Euler

$$\begin{aligned} x_{n+1} &= x_n + hk_1, \\ (I - hJ)k_1 &= f(x_n). \end{aligned}$$

- Order one.
- Stability function as for implicit Euler.

Rosenbrock method of second order

$$\begin{aligned} x_{n+1} &= x_n + \frac{3}{2}hk_1 + \frac{1}{2}hk_2, \\ (I - \gamma hJ)k_1 &= f(x_n), \\ (I - \gamma hJ)k_2 &= f(x_n + hk_1) - 2k_1, \end{aligned}$$

- γ appears in the stability function.
  - $\gamma \geq 1/4$, A-stable.
  - $\gamma = 1 + 1/\sqrt{2}$, L-stable.

[Verwer et al. '99] studied the second order method applied to atmospheric dispersion problems.

Rosenbrock schemes

Schemes

Non-autonomous case

Linearly implicit Euler

$$\begin{aligned} x_{n+1} &= x_n + hk_1, \\ (I - hJ)k_1 &= f(x_n) + hf_t. \end{aligned}$$

- Order one.
- Stability function as for implicit Euler.

$J = \frac{\partial f}{\partial x}(t_n, x_n),\ f_t = \frac{\partial f}{\partial t}(t_n, x_n).$

Rosenbrock method of second order

$$\begin{aligned} x_{n+1} &= x_n + \frac{3}{2}hk_1 + \frac{1}{2}hk_2, \\ (I - \gamma hJ)k_1 &= f(x_n) + \gamma hf_t, \\ (I - \gamma hJ)k_2 &= f(x_n + hk_1) - 2k_1 - \gamma hf_t, \end{aligned}$$

- γ appears in the stability function.
  - $\gamma \geq 1/4$, A-stable.
  - $\gamma = 1 + 1/\sqrt{2}$, L-stable.

Application to DREs

Linear implicit Euler

We consider DREs of the form

$$\begin{array}{rcl} \dot{X}(t) & = & Q(t) + A^T(t)X(t) + X(t)A(t) - X(t)R(t)X(t) \equiv F(t, X(t)) \\ X(t_0) & = & X_0, \qquad t_0 \leq t \leq T. \end{array}$$

The (Frechét) derivative of F at X_k is given by the Lyapunov operator

$$\mathcal{F}^{'}(X_k) : U \rightarrow (A_k - S_k X_k)^T U + U(A_k - S_k X_k),$$

where $X_k \approx X(t_k)$, $A_k = A(t_k)$, $S_k = S(t_k)$ and $U \in \mathbb{R}^{n \times n}$.

Linear implicit Euler method

$$\begin{array}{rcl} X_{k+1} & = & X_k + hK_1, \\ \bar{A}_k^T K_1 + K_1 \bar{A}_k & = & -F(X_k). \end{array}$$

where $\bar{A}_k = h(A_k - S_k X_k) - \frac{1}{2} I$.

Application to DREs

Linear implicit Euler

We consider DREs of the form

$$\begin{array}{rcl} \dot{X}(t) & = & Q(t) + A^T(t)X(t) + X(t)A(t) - X(t)R(t)X(t) \equiv F(t, X(t)) \\ X(t_0) & = & X_0, \qquad t_0 \leq t \leq T. \end{array}$$

The (Frechét) derivative of F at X_k is given by the Lyapunov operator

$$\mathcal{F}^{'}(X_k) : U \rightarrow (A_k - S_k X_k)^T U + U(A_k - S_k X_k),$$

where $X_k \approx X(t_k)$, $A_k = A(t_k)$, $S_k = S(t_k)$ and $U \in \mathbb{R}^{n \times n}$.

Linear implicit Euler method

$$\begin{array}{rcl} X_{k+1} & = & X_k + hK_1, \\ \bar{A}_k^T K_1 + K_1 \bar{A}_k & = & -F(X_k) - hF_{t_k}. \end{array}$$

where $\bar{A}_k = h(A_k - S_k X_k) - \frac{1}{2} I$ and $F_{t_k} \approx \frac{\partial F}{\partial t}(t_k, X(t_k))$.

Application to DREs

Linear implicit Euler (one step)

If we write $F(X_k)$ as

$$\left(A_k - S_k X_k - \tfrac{1}{2h} I\right)^T X_k + X_k \left(A_k - S_k X_k - \tfrac{1}{2h} I\right) \\ + Q_k + X_k S_k X_k + \tfrac{1}{h} X_k,$$

and denoting $\tilde{A}_k = A_k - S_k X_k - \frac{1}{2h} I$, then we can re-write the linear implicit Euler method as

linear implicit Euler method (one step)

$$\tilde{A}_k^T X_{k+1} + X_{k+1} \tilde{A}_k = -Q_k - X_k S_k X_k - \frac{1}{h} X_k. \qquad (1)$$

Application to DREs

Linear implicit Euler (one step)

If we write $F(X_k)$ as

$$\begin{array}{c}\left(A_k - S_k X_k - \frac{1}{2h} I\right)^T X_k + X_k \left(A_k - S_k X_k - \frac{1}{2h} I\right) \\ + Q_k + X_k S_k X_k + \frac{1}{h} X_k,\end{array}$$

and denoting $\tilde{A}_k = A_k - S_k X_k - \frac{1}{2h} I$, then we can re-write the linear implicit Euler method as

linear implicit Euler method (one step) non-autonomous

$$\tilde{A}_k^T X_{k+1} + X_{k+1} \tilde{A}_k = -Q_k - X_k S_k X_k - \frac{1}{h} X_k - h F_{t_k}. \qquad (2)$$

Application to DREs

Rosenbrock method of second order

Rosenbrock method of second order

$$
\begin{aligned}
X_{k+1} &= X_k + \frac{3}{2} h K_1 + \frac{1}{2} h K_2, \\
\hat{A}_k^T K_1 + K_1 \hat{A}_k &= -F(X_k), \\
\hat{A}_k^T K_2 + K_2 \hat{A}_k &= -F(X_k + hK_1) + 2K_1.
\end{aligned}
$$

where:

$$\hat{A}_k = \gamma h (A_k - S_k X_k) - \frac{1}{2} I, \qquad t_{k+1} = t_k + h$$

- Suitable for changing the step size.

Application to DREs

Rosenbrock method of second order

If $\gamma = 1$

Rosenbrock method of second order

$$\begin{aligned} X_{k+1} &= X_k + \frac{3}{2} h K_1 + \frac{1}{2} h K_2, \\ \hat{A}_k^T K_1 + K_1 \hat{A}_k &= -F(X_k), \\ \hat{A}_k^T K_2 + K_2 \hat{A}_k &= -h^2 K_1 S K_1 - K_1. \end{aligned}$$

where:

$$\hat{A}_k = h(A_k - S_k X_k) - \frac{1}{2} I, \qquad t_{k+1} = t_k + h$$

- A-stability is achieved.

Application to DREs

Rosenbrock method of second order

Rosenbrock method of second order non-autonomous

$$\begin{aligned} X_{k+1} &= X_k + \frac{3}{2}hK_1 + \frac{1}{2}hK_2, \\ \hat{A}_k^T K_1 + K_1\hat{A}_k &= -F(X_k) - \gamma h F_{t_k}, \\ \hat{A}_k^T K_2 + K_2\hat{A}_k &= -F(t_{t+1}, X_k + hK_1) + 2K_1 + \gamma h F_{t_k}. \end{aligned}$$

where:

$$\hat{A}_k = \gamma h(A_k - S_k X_k) - \frac{1}{2}I, \qquad t_{k+1} = t_k + h$$

Application to DREs

Large-scale DREs arising in optimal control of PDEs

Consider a semi-discrete LQR problem for a parabolic PDE

$$\begin{aligned} \dot{x}_h &= A_h x_h + B_h u, \\ y_h &= C_h x_h \end{aligned}$$

with cost function

$$J(u) = \int_0^{T_f} < y_h, Q_h y_h > + < u, Ru > dt + < x_{T_f}, G x_{T_f} >,$$

then u is given in feedback form as

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

where X_h is the solution of

Differential Riccati Equation(DRE)

$$\dot{X} = -(C_h Q_h C_h + A_h^T X + X A_h - X B_h R^{-1} B_h^T X), \qquad X(T_f) = G.$$

Application to DREs

Large-scale DREs arising in optimal control of PDEs

Consider a semi-discrete LQR problem for a parabolic PDE

$$\begin{aligned} M\dot{x}_h &= A_h x_h + B_h u, \\ y_h &= C_h x_h \end{aligned}$$

with cost function

$$J(u) = \int_0^{T_f} < y_h, Q_h y_h > + < u, Ru > dt + < x_{T_f}, G x_{T_f} >,$$

then u is given in feedback form as

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

where X_h is the solution of

Differential Riccati Equation(DRE)

$$\dot{X} = -(C_h Q_h C_h + A_h^T X M + M X A_h - M X B_h R^{-1} B_h^T X M), \qquad X(T_f) = G.$$

Application to DREs

Large-scale DREs arising in optimal control of PDEs

General form for $A, R = R^T, Q = Q^T \in \mathbb{R}^{n\times n}$ given and $X \in \mathbb{R}^{n\times n}$ unknown:

DRE

$$\begin{aligned} \dot{X}(t) &= -(Q + A^T X(t) + P(t)A - X(t)RX(t)), \\ X(T_f) &= G. \end{aligned}$$

Here, control-theoretic assumptions ensure existence and uniqueness of the solution.

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

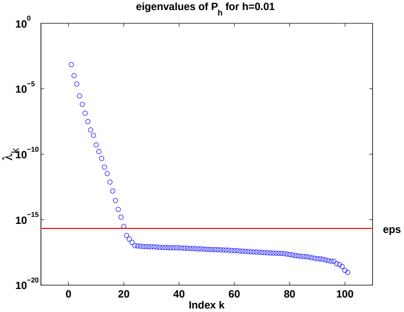
- $n = 10^3 - 10^6$ ($\Longrightarrow$ $10^6 - 10^{12}$ unknowns!),
- A has sparse representation ($A = -M^{-1}L$),
- R, Q low-rank with
  - $R = BB^T, \quad B \in \mathbb{R}^{n\times m}$, $m \ll n$,
  - $Q = C^T C, \quad C \in \mathbb{R}^{p\times n}$, $p \ll n$.

Large-scale DREs arising in optimal control of PDEs

Low-rank approximation

Consider the spectrum of ARE solution

- Linear 1D heat equation with point control.
- $\Omega = [0, 1]$.
- FEM discretization using linear B-splines.
- h=0.01.



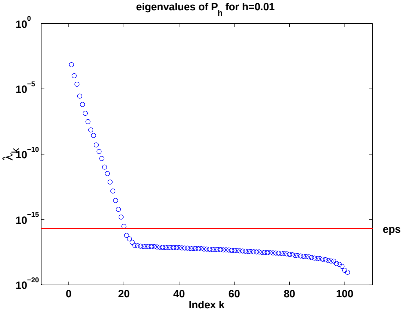
$$P = P^T \geq 0 \Longrightarrow P = ZZ^T = \sum_{k=1}^{n} \lambda_k z_k z_k^T \approx Z^{(r)}(Z^{(r)})^T = \sum_{k=1}^{r} \lambda_k z_k z_k^T.$$

Large-scale DREs arising in optimal control of PDEs

Low-rank approximation

Consider the spectrum of ARE solution

- Linear 1D heat equation with point control.
- $\Omega = [0, 1]$.
- FEM discretization using linear B-splines.
- h=0.01.



$$P = P^T \geq 0 \Longrightarrow P = ZZ^T = \sum_{k=1}^{n} \lambda_k z_k z_k^T \approx Z^{(r)}(Z^{(r)})^T = \sum_{k=1}^{r} \lambda_k z_k z_k^T.$$

Low-rank Rosenbrock methods

ADI Method for Lyapunov equations

Lyapunov equation

$$F^T P + PF = -GG^T$$

ADI iteration for the Lyapunov equation

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

[Penzl 1999; Li/White 2002] rewrite to iterate on the low rank Cholesky factors Q_j of P_j to exploit $\mathrm{rank}(P_j) \ll n$ ($G \in \mathbb{R}^{n \times p}$, $p \ll n$)

Low-rank Rosenbrock methods

Schemes

Let us assume,

$$\begin{array}{rcll} Q & = & C^T C, & C \in \mathbb{R}^{p\times n}, \\ S & = & BB^T, & B \in \mathbb{R}^{n\times m}, \\ X_k & = & Z_k Z_k^T, & Z_k \in \mathbb{R}^{n\times z_k}. \end{array}$$

with p, m, $z_k \ll n$. If we denote $N_k = [\, C^T \;\; Z_k(Z_k^T B) \;\; \sqrt{h^{-1}} Z_k \,]$, then the linearly implicit Euler method

$$\tilde{A}_0^T X_{k+1} + X_{k+1}\tilde{A}_0 = -N_k N_k^T$$

where $\tilde{A}_0 = A - B(Z_k(Z_k^T B))^T - \frac{1}{2h} I$.

Low-rank Rosenbrock methods

Schemes

Let us assume,

$$\begin{array}{rcll} Q & = & C^T C, & C \in \mathbb{R}^{p\times n}, \\ S & = & BB^T, & B \in \mathbb{R}^{n\times m}, \\ X_k & = & Z_k Z_k^T, & Z_k \in \mathbb{R}^{n\times z_k}. \end{array}$$

with p, m, $z_k \ll n$. If we denote $N_k = [\, C^T \;\; Z_k(Z_k^T B) \;\; \sqrt{h^{-1}} Z_k \,]$, then the linearly implicit Euler method

$$\tilde{A}_0^T X_{k+1} + X_{k+1}\tilde{A}_0 = -N_k N_k^T$$

where $\tilde{A}_0 = A - B(Z_k(Z_k^T B))^T - \frac{1}{2h} I$.

- $\mathrm{rank}(N_k) \leq p + m + z_k \ll n$ → LRF-ADI iteration.

Low-rank Rosenbrock methods

Schemes

Rosenbrock method of second order

$$\begin{aligned} X_{k+1} &= X_k + \frac{3}{2}hK_1 + \frac{1}{2}hK_2, \\ \hat{A}_k^T K_1 + K_1 \hat{A}_k &= -F(X_k), \\ \hat{A}_k^T K_2 + K_2 \hat{A}_k &= -F(X_k + hK_1) + 2K_1. \end{aligned}$$

where:

$$\hat{A}_k = \gamma h(A_k - S_k X_k) - \frac{1}{2}I, \qquad t_{k+1} = t_k + h$$

- Idea: write right hand side of Lyapunov equations as low rank matrix product.

Low-rank Rosenbrock methods

Schemes

Rosenbrock method of second order

$$\begin{aligned} X_{k+1} &= X_k + \frac{3}{2}hK_1 + \frac{1}{2}hK_2, \\ \hat{A}_k^T K_1 + K_1 \hat{A}_k &= -F(X_k), \\ \hat{A}_k^T K_2 + K_2 \hat{A}_k &= -F(X_k + hK_1) + 2K_1. \end{aligned}$$

where:

$$\hat{A}_k = \gamma h(A_k - S_k X_k) - \frac{1}{2}I, \qquad t_{k+1} = t_k + h$$

- Idea: write right hand side of Lyapunov equations as low rank matrix product.

Low-rank Rosenbrock methods

Schemes

For $k = 0$

$$A^T Z_0 Z_0^T + Z_0 Z_0^T A = (A^T Z_0 + Z_0)(A^T Z_0 + Z_0)^T - [\, A^T Z_0 \;\; Z_0 \,][\, A^T Z_0 \;\; Z_0 \,]^T.$$

Let $L_0 := A^T Z_0 + Z_0$, $M_0 := [\, A^T Z_0 \;\; Z_0 \,]$, and $W_0 := Z_0(Z_0^T B)$, denoting $N_0 := [\, C^T \;\; L_0 \,]$, $U_0 := [\, M_0 \;\; W_0 \,]$, we can split the first Lyapunov equation into

$$\begin{aligned} \hat{A}_0^T \tilde{K}_1 + \tilde{K}_1 \hat{A}_0 &= -U_0 U_0^T && (3) \\ \hat{A}_0^T \hat{K}_1 + \hat{K}_1 \hat{A}_0 &= -N_0 N_0^T && (4) \end{aligned}$$

where $K_1 := \hat{K}_1 - \tilde{K}_1$.

- Low-rank factor of K_1 becomes complex.
- K_1 is probably no longer positive semidefinite.

Low-rank Rosenbrock methods

Schemes

For $k = 0$

$$A^T Z_0 Z_0^T + Z_0 Z_0^T A = (A^T Z_0 + Z_0)(A^T Z_0 + Z_0)^T - [\, A^T Z_0 \;\; Z_0 \,][\, A^T Z_0 \;\; Z_0 \,]^T.$$

Let $L_0 := A^T Z_0 + Z_0$, $M_0 := [\, A^T Z_0 \;\; Z_0 \,]$, and $W_0 := Z_0(Z_0^T B)$, denoting $N_0 := [\, C^T \;\; L_0 \,]$, $U_0 := [\, M_0 \;\; W_0 \,]$, we can split the first Lyapunov equation into

$$\begin{aligned} \hat{A}_0^T \tilde{K}_1 + \tilde{K}_1 \hat{A}_0 &= -U_0 U_0^T \qquad (3) \\ \hat{A}_0^T \hat{K}_1 + \hat{K}_1 \hat{A}_0 &= -N_0 N_0^T \qquad (4) \end{aligned}$$

where $K_1 := \hat{K}_1 - \tilde{K}_1$.

- Low-rank factor of K_1 becomes complex.
- K_1 is probably no longer positive semidefinite.

Low-rank Rosenbrock methods

Schemes

- Idea: work with the low rank factors of $\hat{K}_1$ and $\tilde{K}_1$ in the iteration.
  - Prove $X_k \approx Z_0 Z_0^T + h(\hat{T}_k \hat{T}_k^T - \tilde{T}_k \tilde{T}_k^T)$ for each k.
  - Rewrite $-F(X_k)$, $-F(X_k + hK_1) + 2K_1$ as a matrix product.
- Each Lyapunov equation is split.
  - The resulting Lyapunov equations can be solved simultaneously.
  - Column compression technique maybe needed.
- The special case $\gamma = 1$ simplifies the iteration.
- Our analysis can be extended to s-stage Rosenbrock methods.

F_{t_k} can be represented as a low rank matrix product combination *to* non-autonomous case.

Low-rank Rosenbrock methods

Schemes

- Idea: work with the low rank factors of $\hat{K}_1$ and $\tilde{K}_1$ in the iteration.
  - Prove $X_k \approx Z_0 Z_0^T + h(\hat{T}_k \hat{T}_k^T - \tilde{T}_k \tilde{T}_k^T)$ for each k.
  - Rewrite $-F(X_k)$, $-F(X_k + hK_1) + 2K_1$ as a matrix product.
- Each Lyapunov equation is split.
  - The resulting Lyapunov equations can be solved simultaneously.
  - Column compression technique maybe needed.
- The special case $\gamma = 1$ simplifies the iteration.
- Our analysis can be extended to s-stage Rosenbrock methods.

F_{t_k} can be represented as a low rank matrix product combination *to* non-autonomous case.

Numerical examples

Small-scale

Example 1. [Choi/Laub 1990]

Let us now consider the following symmetric DRE of size n,

$$\begin{cases} \dot{X}(t) = -X^2(t) + k^2 I_n \\ X(t_0) = X_0 \qquad t_0 \leq t \leq T \end{cases}$$

If X_0 is diagonalizable, i.e, $X_0 = S\Lambda S^{-1}$ where Λ=diag$[\lambda_i]$, then its analytic solution is:

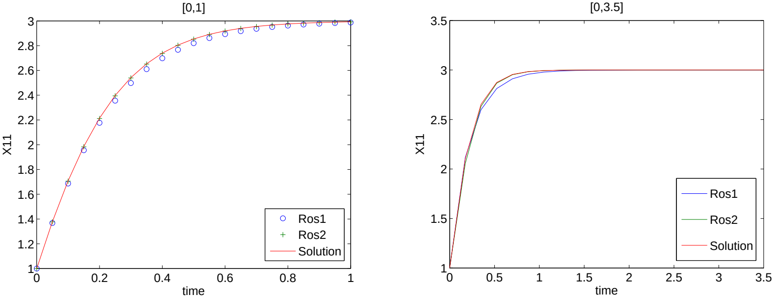
$$X(t) = S\mathrm{diag}\left[\frac{k \sinh kt + \lambda_i \cosh kt}{\cosh kt + \frac{\lambda_i}{k} \sinh kt}\right] S^{-1}.$$

Here, we choose

$$X_0 = I_n, \quad k = 3, \quad n = 60.$$

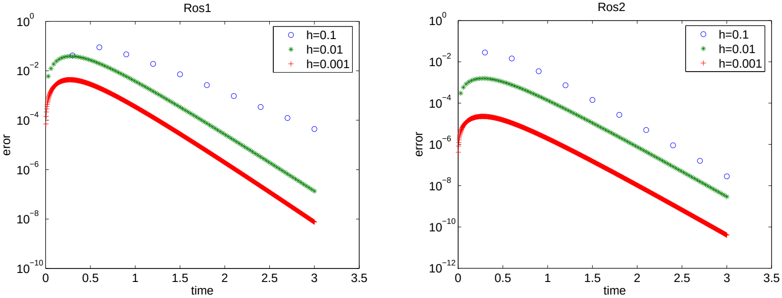
Numerical examples

Small-scale



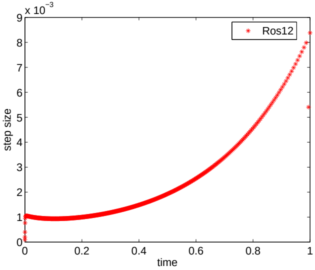
Numerical examples

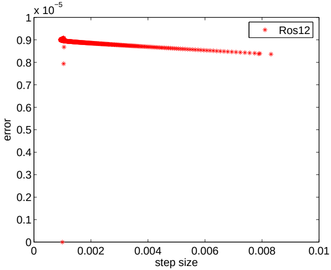
Small-scale

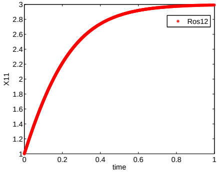


Numerical examples

Small-scale





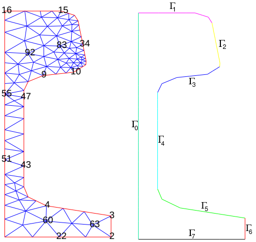


Numerical examples

Large-scale

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

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

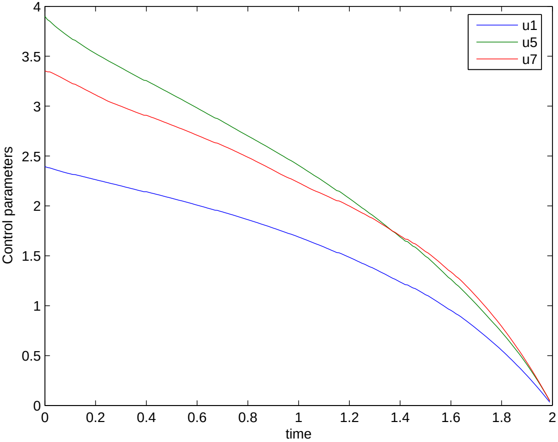


- FEM discretization, different models for initial mesh ($n = 371$),
  1, 2, 3, 4 steps of mesh refinement ⇒
  $n = 1357, 5177, 20209, 79841$.

Source: Physical model: courtesy of Mannesmann/Demag.
Math. model: Tröltzsch/Unger 1999/2001, Penzl 1999, S. 2003.

Numerical examples

Large-scale



Conclusion and Outlook

Conclusions

- Rosenbrock schemes allow an efficient implementation for large-scale DREs arising in optimal control for PDEs.

Outlook

- Computation of the feedback matrices directly.
- Parallel implementation.

The End

Thank you for your attention!