Numerical Analysis and Scientific Computation
with Applications (NASCA13)
Calais, June 24–26, 2013

Numerical Solution of Linear and Nonlinear Matrix Equations Arising in Stochastic and Bilinear Control Theory

Peter Benner Tobias Breiten

Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory
Magdeburg, Germany

Overview

Introduction

Dynamical Systems

Dynamical Systems

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) = f(t, x(t;p), u(t), p), \quad x(t_0) = x_0, & \text{(a)} \\ y(t;p) = g(t, x(t;p), u(t), p) & \text{(b)} \end{cases}$$

with

- (generalized) states $x(t;p) \in \mathbb{R}^n$ ($E \in \mathbb{R}^{n \times n}$),
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t;p) \in \mathbb{R}^q$, (b) is called output equation,
- $p \in \Omega \subset \mathbb{R}^d$ is a parameter vector, Ω is bounded.

Applications:

- Repeated simulation for varying material or geometry parameters, boundary conditions,
- control, optimization and design.

Linear Parametric Systems

Linear, time-invariant (parametric) systems

$$
\begin{aligned}
E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \quad && A(p), E(p) \in \mathbb{R}^{n\times n},\\
y(t;p) &= C(p)x(t;p), && B(p) \in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}.
\end{aligned}
$$

Linear Parametric Systems

Linear, time-invariant (parametric) systems

$$\begin{array}{rcll} E(p)\dot{x}(t;p) & = & A(p)x(t;p) + B(p)u(t), & A(p), E(p) \in \mathbb{R}^{n\times n}, \\ y(t;p) & = & C(p)x(t;p), & B(p) \in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}. \end{array}$$

Laplace Transformation / Frequency Domain

Application of Laplace transformation $\quad (x(t;p) \mapsto x(s;p),\ \dot{x}(t;p) \mapsto sx(s;p))$ to linear system with $x(0) = 0$:

$$sE(p)x(s;p) = A(p)x(s;p) + B(p)u(s), \quad y(s;p) = C(p)x(s;p),$$

yields I/O-relation in frequency domain:

$$y(s;p) = \Big(\underbrace{C(p)(sE(p) - A(p))^{-1}B(p)}_{=:H(s;p)}\Big)u(s).$$

$H(s;p)$ is the parameter-dependent transfer function of $\Sigma(p)$.

Linear Parametric Systems

Linear, time-invariant (parametric) systems

$$\begin{aligned} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \quad && A(p), E(p) \in \mathbb{R}^{n\times n}, \\ y(t;p) &= C(p)x(t;p), && B(p) \in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}. \end{aligned}$$

Laplace Transformation / Frequency Domain

Application of Laplace transformation $\quad (x(t;p) \mapsto x(s;p),\ \dot{x}(t;p) \mapsto sx(s;p))$ to linear system with $x(0) = 0$:

$$sE(p)x(s;p) = A(p)x(s;p) + B(p)u(s), \quad y(s;p) = C(p)x(s;p),$$

yields I/O-relation in frequency domain:

$$y(s;p) = \Big(\underbrace{C(p)(sE(p) - A(p))^{-1}B(p)}_{=:H(s;p)} \Big) u(s).$$

$H(s;p)$ is the parameter-dependent transfer function of $\Sigma(p)$.

Goal: Fast evaluation of mapping $(u,p) \ \rightarrow \ y(s;p)$.

Introduction

Model Order Reduction (MOR) Problem

Problem

Approximate the dynamical system

$$\begin{aligned} E(p)\dot{x} &= A(p)x + B(p)u, & E(p), A(p) &\in \mathbb{R}^{n\times n}, \\ y &= C(p)x, & B(p) &\in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}, \end{aligned}$$

by reduced-order system

$$\begin{aligned} \hat{E}(p)\dot{\hat{x}} &= \hat{A}(p)\hat{x} + \hat{B}(p)u, & \hat{E}(p), \hat{A}(p) &\in \mathbb{R}^{r\times r}, \\ \hat{y} &= \hat{C}(p)\hat{x}, & \hat{B}(p) &\in \mathbb{R}^{r\times m}, \hat{C}(p) \in \mathbb{R}^{q\times r}, \end{aligned}$$

of order $r \ll n$, such that

$$\|y - \hat{y}\| = \|Hu - \hat{H}u\| \leq \|H - \hat{H}\| \cdot \|u\| < \text{tolerance} \cdot \|u\| \quad \forall\, p \in \Omega.$$

Introduction

Model Order Reduction (MOR) Problem

Problem

Approximate the dynamical system

$$\begin{array}{rcll} E(p)\dot{x} & = & A(p)x + B(p)u, & E(p), A(p) \in \mathbb{R}^{n\times n}, \\ y & = & C(p)x, & B(p) \in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}, \end{array}$$

by reduced-order system

$$\begin{array}{rcll} \hat{E}(p)\dot{\hat{x}} & = & \hat{A}(p)\hat{x} + \hat{B}(p)u, & \hat{E}(p), \hat{A}(p) \in \mathbb{R}^{r\times r}, \\ \hat{y} & = & \hat{C}(p)\hat{x}, & \hat{B}(p) \in \mathbb{R}^{r\times m}, \hat{C}(p) \in \mathbb{R}^{q\times r}, \end{array}$$

of order $r \ll n$, such that

$$\|y - \hat{y}\| = \|Hu - \hat{H}u\| \leq \|H - \hat{H}\| \cdot \|u\| < \text{tolerance} \cdot \|u\| \quad \forall\ p \in \Omega.$$

$\Longrightarrow$ Approximation problem: $\min\limits_{\text{order}\,(\hat{H})\leq r} \|H - \hat{H}\|.$

Introduction to Model Order Reduction

Linear Parametric Systems — An Alternative Interpretation

Consider bilinear control systems:

$$\Sigma: \quad \begin{cases} \dot{x}(t) = Ax(t) + \sum_{i=1}^{m} N_i x(t) u_i(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

where $A, N_i \in \mathbb{R}^{n \times n}, \; B \in \mathbb{R}^{n \times m}, \; C \in \mathbb{R}^{q \times n}$.

Introduction to Model Order Reduction

Linear Parametric Systems — An Alternative Interpretation

Consider bilinear control systems:

$$\Sigma : \quad \begin{cases} \dot{x}(t) = Ax(t) + \sum_{i=1}^{m} N_i x(t) u_i(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

where $A, N_i \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{q \times n}$.

Key Observation: Regarding parameter dependencies as additional inputs, a linear parametric system

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m_p} a_i(p) A_i x(t) + B_0 u_0(t), \quad y(t) = Cx(t)$$

with $B_0 \in \mathbb{R}^{n \times m_0}$ can be interpreted as bilinear system (with $N_i \equiv A_i$):

$$u(t) := \begin{bmatrix} a_1(p) & \dots & a_{m_p}(p) & u_0(t) \end{bmatrix}^T,$$

$$B := \begin{bmatrix} \mathbf{0} & \dots & \mathbf{0} & B_0 \end{bmatrix} \in \mathbb{R}^{n \times m}, \quad m = m_p + m_0.$$

Balanced truncation for linear systems

Idea (for simplicity, $E = I_n$)

- $\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,

  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

  $$AP + PA^T + BB^T = 0, \qquad A^T Q + QA + C^T C = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

Balanced truncation for linear systems

Idea (for simplicity, $E = I_n$)

- $\Sigma: \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,

  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

  $$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.
- $\{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.

Balanced truncation for linear systems

Idea (for simplicity, $E = I_n$)

- $\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,
  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

  $$AP + PA^T + BB^T = 0, \qquad A^T Q + QA + C^T C = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.
- $\{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization of the system via state-space transformation

  $$\begin{aligned} \mathcal{T} : (A, B, C) \quad &\mapsto \quad (TAT^{-1}, TB, CT^{-1}) \\ &= \quad \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix} \right). \end{aligned}$$

Balanced truncation for linear systems

Idea (for simplicity, $E = I_n$)

- $\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,
  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

$$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$.
- $\{\sigma_1, \dots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization of the system via state-space transformation

$$\begin{aligned} \mathcal{T} : (A, B, C) &\mapsto (TAT^{-1}, TB, CT^{-1}) \\ &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix} \right). \end{aligned}$$

- Truncation $\rightsquigarrow (\hat{A}, \hat{B}, \hat{C}) = (A_{11}, B_1, C_1)$.

Balanced truncation for linear systems

Motivation:

HSV are system invariants: they are preserved under $\mathcal{T}$ and determine the energy transfer given by the Hankel map

$$\mathcal{H} : L_2(-\infty, 0) \mapsto L_2(0, \infty) : u_- \mapsto y_+.$$

"functional analyst's point of view"

Balanced truncation for linear systems

Motivation:

HSV are system invariants: they are preserved under $\mathcal{T}$ and determine the energy transfer given by the Hankel map

$$\mathcal{H} : L_2(-\infty, 0) \mapsto L_2(0, \infty) : u_- \mapsto y_+.$$

"functional analyst's point of view"

In balanced coordinates, energy transfer from u_- to y_+ is

$$E := \sup_{\substack{u \in L_2(-\infty,0] \\ x(0)=x_0}} \frac{\int_0^\infty y(t)^T y(t)\, dt}{\int_{-\infty}^0 u(t)^T u(t)\, dt} = \frac{1}{\|x_0\|_2} \sum_{j=1}^n \sigma_j^2 x_{0,j}^2.$$

"engineer's point of view"

Balanced truncation for linear systems

Motivation:

HSV are system invariants: they are preserved under $\mathcal{T}$ and determine the energy transfer given by the Hankel map

$$\mathcal{H} : L_2(-\infty, 0) \mapsto L_2(0, \infty) : u_- \mapsto y_+.$$

"functional analyst's point of view"

In balanced coordinates, energy transfer from u_- to y_+ is

$$E := \sup_{\substack{u \in L_2(-\infty,0] \\ x(0)=x_0}} \frac{\int_0^\infty y(t)^T y(t)\, dt}{\int_{-\infty}^0 u(t)^T u(t)\, dt} = \frac{1}{\|x_0\|_2} \sum_{j=1}^n \sigma_j^2 x_{0,j}^2.$$

"engineer's point of view" $\Longrightarrow$ Truncate states corresponding to "small" HSVs

$\Longrightarrow$ analogy to best approximation via SVD, therefore balancing-related methods are sometimes called SVD methods.

Balanced truncation for linear systems

Implementation: SR Method

1. Compute (Cholesky) factors of the solutions of the Lyapunov equations,

$$P = S^T S, \quad Q = R^T R.$$

Balanced truncation for linear systems

Implementation: SR Method

1. Compute (Cholesky) factors of the solutions of the Lyapunov equations,

$$P = S^T S, \quad Q = R^T R.$$

2. Compute SVD

$$SR^T = [\, U_1,\ U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$$

Balanced truncation for linear systems

Implementation: SR Method

1. Compute (Cholesky) factors of the solutions of the Lyapunov equations,
$$P = S^T S, \quad Q = R^T R.$$
2. Compute SVD
$$SR^T = [\, U_1, \; U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$$
3. Set
$$W = R^T V_1 \Sigma_1^{-1/2}, \qquad V = S^T U_1 \Sigma_1^{-1/2}.$$
4. Reduced model is $(W^T AV, W^T B, CV)$.

Balanced truncation for linear systems

Implementation: SR Method

1. Compute (Cholesky) factors of the solutions of the Lyapunov equations,
$$P = S^T S, \quad Q = R^T R.$$
2. Compute SVD
$$SR^T = [\, U_1, \, U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$$
3. Set
$$W = R^T V_1 \Sigma_1^{-1/2}, \qquad V = S^T U_1 \Sigma_1^{-1/2}.$$
4. Reduced model is $(W^T AV, W^T B, CV)$.

Note: $T := \Sigma^{-\frac{1}{2}} V^T R$ yields balancing state-space transformation with $T^{-1} = S^T U \Sigma^{-\frac{1}{2}}$, so that $W^T = T(1:r,:)$ and $V = T^{-1}(:,1:r)$.

Balanced truncation for linear systems

Properties:

- Reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$.

Balanced truncation for linear systems

Properties:

- Reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$.
- Adaptive choice of r via computable error bound:

$$\|y - \hat{y}\|_2 \leq \left(2 \sum_{k=r+1}^{n} \sigma_k\right) \|u\|_2.$$

LQG Balanced Truncation

Instead of system Gramians P, Q, use solutions of algebraic Riccati equations (AREs)

$$\begin{aligned} 0 &= AP + PA^T - PC^TCP + BB^T, \\ 0 &= A^TQ + QA - QBB^TQ + C^TC, \end{aligned}$$

related to linear-quadratic Gaussian (LQG) control design.

Properties:

- Applicable to unstable systems.
- When factorizations $P = S^TS,\ Q = R^TR$ are available, construction of reduced-order model exactly as in SR method for balanced truncation.
- Error bound: $"\|H - \hat{H}\|_{L_\infty}" \leq 2 \sum_{j=r+1}^{n} \frac{\gamma_j}{\sqrt{1+\gamma_j^2}}$, where

  $\{\gamma_1, \ldots, \gamma_n\} = \Lambda(PQ)^{\frac{1}{2}}$, $\gamma_1 \geq \ldots \gamma_n \geq 0$.

Summary of Introduction

Balancing-based MOR of linear systems requires the efficient numerical solution of either linear or nonlinear matrix equations:

Balanced truncation: Lyapunov equations

$$\begin{aligned} 0 &= AP + PA^T + BB^T &= \mathcal{L}(P), \\ 0 &= A^T Q + QA + C^T C &= \mathcal{L}^*(Q). \end{aligned}$$

LQG Balanced truncation: algebraic Riccati equations

$$\begin{aligned} 0 &= AP + PA^T - PC^T CP + BB^T &= \mathcal{L}(P) - PC^T CP, \\ 0 &= A^T Q + QA - QBB^T Q + C^T C &= \mathcal{L}^*(Q) - QBB^T Q. \end{aligned}$$

Numerous mature methods exist, e.g.,

- for Lyapunov equations: (rational) Krylov subspace methods, low-rank ADI, Riemannian optimization, ...
- for AREs: (rational) Krylov subspace methods, Newton-ADI, Chandrasekhar iteration, ...

Bilinear Lyapunov Equations

Bilinear Control Systems — Theory and Background

Bilinear control systems:

$$\Sigma : \quad \begin{cases} \dot{x}(t) = Ax(t) + \sum_{i=1}^{m} N_i x(t) u_i(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

where $A, N_i \in \mathbb{R}^{n \times n}, \; B \in \mathbb{R}^{n \times m}, \; C \in \mathbb{R}^{q \times n}$.

Properties:

- Approximation of (weakly) nonlinear systems by Carleman linearization yields bilinear systems.
- Appear naturally in boundary control problems, control via coefficients of PDEs, Fokker-Planck equations, ...
- Due to the close relation to linear systems, a lot of successful concepts can be extended, e.g. transfer functions, Gramians, Lyapunov equations, ...
- Linear stochastic control systems possess an equivalent structure and can be treated alike [B./Damm '11].

Bilinear Lyapunov Equations

Balanced truncation for bilinear systems

The concept of balanced truncation can be generalized to the case of bilinear systems, where we need the solutions of the generalized Lyapunov equations:

$$AP + PA^T + \sum_{i=1}^{m} N_i P A_i^T + BB^T = 0,$$
$$A^T Q + QA^T + \sum_{i=1}^{m} N_i^T Q A_i + C^T C = 0.$$

- Due to its approximation quality, first method of choice for medium-size systems.
- These equations also appear for stochastic control systems, see [B./Damm '11].
- For an iterative full-rank solver, see [Damm '08].

Bilinear Lyapunov Equations

Some basic facts and assumptions

$$AX + XA^T + \sum_{i=1}^{m} N_i X N_i^T + BB^T = 0. \quad (1)$$

- Need a positive semi-definite symmetric solution X.

Bilinear Lyapunov Equations

Some basic facts and assumptions

$$AX + XA^T + \sum_{i=1}^{m} N_i X N_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable ($\Lambda(A) \subset \mathbb{C}^-$); this is not sufficient here: (1) is equivalent to

$$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} N_i \otimes N_i \right) \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$

  One sufficient condition for stable A is smallness of N_i (related to stability radius of $\mathcal{A}$)

  $\rightsquigarrow$ bounded-input bounded-output (BIBO) stability of bilinear systems.

  This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.

Bilinear Lyapunov Equations

Some basic facts and assumptions

$$AX + XA^T + \sum_{i=1}^{m} N_i X N_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable ($\Lambda(A) \subset \mathbb{C}^-$); this is not sufficient here: (1) is equivalent to

  $$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} N_i \otimes N_i \right) \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$

  One sufficient condition for stable A is smallness of N_i (related to stability radius of $\mathcal{A}$)

  $\rightsquigarrow$ bounded-input bounded-output (BIBO) stability of bilinear systems.

  This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.
- Want: solution methods for large scale problems, i.e., only matrix-matrix multiplication with A, N_j, solves with (shifted) A allowed!

Bilinear Lyapunov Equations

Some basic facts and assumptions

$$AX + XA^T + \sum_{i=1}^{m} N_i X N_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable ($\Lambda(A) \subset \mathbb{C}^-$); this is not sufficient here: (1) is equivalent to
  $$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} N_i \otimes N_i \right) \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$
  One sufficient condition for stable A is smallness of N_i (related to stability radius of $\mathcal{A}$)
  $\rightsquigarrow$ bounded-input bounded-output (BIBO) stability of bilinear systems.
  This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.
- Want: solution methods for large scale problems, i.e., only matrix-matrix multiplication with A, N_j, solves with (shifted) A allowed!
- Requires to compute data-sparse approximation to generally dense X; here: $X \approx ZZ^T$ with $Z \in \mathbb{R}^{n \times n_Z}$, $n_Z \ll n$!

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0\ ?$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0 \text{ ?}$$

Standard Lyapunov case: [Grasedyck '04]

$$AX + XA^T + BB^T = 0 \iff \underbrace{(I_n \otimes A + A \otimes I_n)}_{=:\mathcal{A}} \operatorname{vec}(X) = -\operatorname{vec}(BB^T).$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Standard Lyapunov case: [Grasedyck '04]

$$AX + XA^T + BB^T = 0 \iff \underbrace{(I_n \otimes A + A \otimes I_n)}_{=:\mathcal{A}} \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$

Apply

$$M^{-1} = -\int_0^\infty \exp(tM)\mathrm{d}t$$

to $\mathcal{A}$ and approximate the integral via (sinc) quadrature $\Rightarrow$

$$\mathcal{A}^{-1} \approx -\sum_{i=-k}^{k} \omega_i \exp(t_k \mathcal{A}),$$

with error $\sim \exp(-\sqrt{k})$ ($\exp(-k)$ if $A = A^T$), then an approximate Lyapunov solution is given by

$$\mathrm{vec}(X) \approx \mathrm{vec}(X_k) = \sum_{i=-k}^{k} \omega_i \exp(t_i \mathcal{A}) \,\mathrm{vec}(BB^T).$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Standard Lyapunov case: [Grasedyck '04]

$$AX + XA^T + BB^T = 0 \iff \underbrace{(I_n \otimes A + A \otimes I_n)}_{=:\mathcal{A}} \operatorname{vec}(X) = -\operatorname{vec}(BB^T).$$

$$\operatorname{vec}(X) \approx \operatorname{vec}(X_k) = \sum_{i=-k}^{k} \omega_i \exp(t_i \mathcal{A}) \operatorname{vec}(BB^T).$$

Now observe that

$$\exp(t_i \mathcal{A}) = \exp\left(t_i(I_n \otimes A + A \otimes I_n)\right) \equiv \exp(t_i A) \otimes \exp(t_i A).$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Standard Lyapunov case: [GRASEDYCK '04]

$$AX + XA^T + BB^T = 0 \iff \underbrace{(I_n \otimes A + A \otimes I_n)}_{=:\mathcal{A}} \operatorname{vec}(X) = -\operatorname{vec}(BB^T).$$

$$\operatorname{vec}(X) \approx \operatorname{vec}(X_k) = \sum_{i=-k}^{k} \omega_i \exp(t_i \mathcal{A}) \operatorname{vec}(BB^T).$$

Now observe that

$$\exp(t_i \mathcal{A}) = \exp\left(t_i (I_n \otimes A + A \otimes I_n)\right) \equiv \exp(t_i A) \otimes \exp(t_i A).$$

Hence,

$$\operatorname{vec}(X_k) = \sum_{i=-k}^{k} \omega_i \left(\exp(t_i A) \otimes \exp(t_i A)\right) \operatorname{vec}(BB^T)$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Standard Lyapunov case: [GRASEDYCK '04]

$$AX + XA^T + BB^T = 0 \iff \underbrace{(I_n \otimes A + A \otimes I_n)}_{=:\mathcal{A}} \operatorname{vec}(X) = -\operatorname{vec}(BB^T).$$

Hence,

$$\begin{aligned} \operatorname{vec}(X_k) &= \sum_{i=-k}^{k} \omega_i \left(\exp(t_i A) \otimes \exp(t_i A)\right) \operatorname{vec}(BB^T) \\ \implies X_k &= \sum_{i=-k}^{k} \omega_i \exp(t_i A) B B^T \exp(t_i A^T) \equiv \sum_{i=-k}^{k} \omega_i B_i B_i^T, \end{aligned}$$

so that $\operatorname{rank}(X_k) \leq (2k+1)m$ with

$$\|X - X_k\|_2 \lesssim \exp(-\sqrt{k}) \qquad (\exp(-k) \text{ for } A = A^T)!$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0 \ ?$$

Problem: in general,

$$\exp\left(t_i(I \otimes A + A \otimes + \sum_{j=1}^{m} N_j \otimes N_j)\right) \neq (\exp(t_i A) \otimes \exp(t_i A)) \exp\left(t_i(\sum_{j=1}^{m} N_j \otimes N_j)\right).$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0 \text{ ?}$$

Assume that $m = 1$ and $N_1 = UV^T$ with $U, V \in \mathbb{R}^{n \times r}$ and consider

$$(\underbrace{I_n \otimes A + A \otimes I_n}_{=\mathcal{A}} + N_1 \otimes N_1)\,\mathrm{vec}(X) = \underbrace{-\,\mathrm{vec}(BB^T)}_{=:y}.$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0 \ ?$$

Assume that $m = 1$ and $N_1 = UV^T$ with $U, V \in \mathbb{R}^{n \times r}$ and consider

$$(\underbrace{I_n \otimes A + A \otimes I_n}_{=\mathcal{A}} + N_1 \otimes N_1)\,\mathrm{vec}(X) = \underbrace{-\mathrm{vec}(BB^T)}_{=:y}.$$

Sherman-Morrison-Woodbury $\Longrightarrow$

$$\begin{aligned}\left(I_r \otimes I_r + (V^T \otimes V^T)\mathcal{A}^{-1}(U \otimes U)\right) w &= (V^T \otimes V^T)\mathcal{A}^{-1} y,\\ \mathcal{A}\,\mathrm{vec}(X) &= y - (U \otimes U) w.\end{aligned}$$

Bilinear Lyapunov Equations

Existence of low-rank approximations

Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0 \ ?$$

Assume that $m = 1$ and $N_1 = UV^T$ with $U, V \in \mathbb{R}^{n\times r}$ and consider

$$(\underbrace{I_n \otimes A + A \otimes I_n}_{=\mathcal{A}} + N_1 \otimes N_1)\,\mathrm{vec}(X) = \underbrace{-\,\mathrm{vec}(BB^T)}_{=:y}.$$

Sherman-Morrison-Woodbury $\Longrightarrow$

$$\begin{aligned}\left(I_r \otimes I_r + (V^T \otimes V^T)\mathcal{A}^{-1}(U \otimes U)\right) w &= (V^T \otimes V^T)\mathcal{A}^{-1} y,\\ \mathcal{A}\,\mathrm{vec}(X) &= y - (U \otimes U)w.\end{aligned}$$

Matrix representation of r.h.s., $-BB^T - U\,\mathrm{vec}^{-1}(w)\,U^T$ has rank $\leq\ r+1$!

$\rightsquigarrow$ Apply results for linear Lyapunov equations with r.h.s of rank $r+1$.

Bilinear Lyapunov Equations

Existence of low-rank approximations

Theorem [B./Breiten 2012]

Assume existence and uniqueness assumption with stable A and $N_j = U_j V_j^T$, with $U_j, V_j \in \mathbb{R}^{n \times r_j}$. Set $r = \sum_{j=1}^m r_j$.
Then the solution X of

$$AX + XA^T + \sum_{j=1}^m N_j X N_j^T + BB^T = 0$$

can be approximated by X_k of rank $(2k+1)(m+r)$, with an error satisfying

$$\|X - X_k\|_2 \lesssim \exp(-\sqrt{k}).$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Let us again consider the generalized Lyapunov equation

$$AP + PA^T + NPN^T + BB^T = 0.$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Let us again consider the generalized Lyapunov equation

$$AP + PA^T + NPN^T + BB^T = 0.$$

For a fixed parameter p, we can rewrite the linear Lyapunov operator as

$$AP + PA^T = \frac{1}{2p}\left((A + pI)P(A + pI)^T - (A - pI)P(A - pI)^T\right)$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Let us again consider the generalized Lyapunov equation

$$AP + PA^T + NPN^T + BB^T = 0.$$

For a fixed parameter p, we can rewrite the linear Lyapunov operator as

$$AP + PA^T = \frac{1}{2p}\left((A+pI)P(A+pI)^T - (A-pI)P(A-pI)^T\right)$$

leading to the fix point iteration [Damm '08]

$$\begin{aligned} P_j = {} & (A-pI)^{-1}(A+pI)P_{j-1}(A+pI)^T(A-pI)^{-T} \\ & + 2p(A-pI)^{-1}(NP_{j-1}N^T + BB^T)(A-pI)^{-T}. \end{aligned}$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Let us again consider the generalized Lyapunov equation

$$AP + PA^T + NPN^T + BB^T = 0.$$

For a fixed parameter p, we can rewrite the linear Lyapunov operator as

$$AP + PA^T = \frac{1}{2p}\left((A+pI)P(A+pI)^T - (A-pI)P(A-pI)^T\right)$$

leading to the fix point iteration [Damm '08]

$$\begin{aligned} P_j &= (A-pI)^{-1}(A+pI)P_{j-1}(A+pI)^T(A-pI)^{-T} \\ &\quad + 2p(A-pI)^{-1}(NP_{j-1}N^T + BB^T)(A-pI)^{-T}. \end{aligned}$$

$P_j \approx Z_j Z_j^T$ $(\mathrm{rank}\,(Z_j) \ll n)$ $\rightsquigarrow$ factored iteration

$$\begin{aligned} Z_j Z_j^T &= (A-pI)^{-1}(A+pI)Z_{j-1}Z_{j-1}^T(A+pI)^T(A-pI)^{-T} \\ &\quad + 2p(A-pI)^{-1}(NZ_{j-1}Z_{j-1}^T N^T + BB^T). \end{aligned}$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Hence, for a given sequence of shift parameters $\{p_1, \ldots, p_q\}$, we can extend the linear ADI iteration as follows:

$$
\begin{aligned}
Z_1 &= \sqrt{2p_1}\,(A - p_1 I)^{-1}\,B,\\
Z_j &= (A - p_j I)^{-1}\left[(A + p_j I)\,Z_{j-1} \quad \sqrt{2p_j}B \quad \sqrt{2p_j}NZ_{j-1}\right], \quad j \leq q.
\end{aligned}
$$

Bilinear Lyapunov Equations

Generalized alternating directions iteration (ADI)

Hence, for a given sequence of shift parameters $\{p_1, \ldots, p_q\}$, we can extend the linear ADI iteration as follows:

$$
\begin{aligned}
Z_1 &= \sqrt{2p_1}\,(A - p_1 I)^{-1}\,B,\\
Z_j &= (A - p_j I)^{-1}\left[(A + p_j I)\,Z_{j-1} \quad \sqrt{2p_j}B \quad \sqrt{2p_j}NZ_{j-1}\right], \quad j \leq q.
\end{aligned}
$$

Problems:

- A and N in general do not commute $\rightsquigarrow$ we have to operate on full preceding subspace Z_{j-1} in each step.
- Rapid increase of $\operatorname{rank}(Z_j)$ $\rightsquigarrow$ perform some kind of column compression.
- Choice of shift parameters? $\rightsquigarrow$ No obvious generalization of minimax problem.
  Here, we will use shifts minimizing a certain $\mathcal{H}_2$-optimization problem, see [B./Breiten '11].

Generalized alternating directions iteration (ADI)

Numerical Example: A Heat Transfer Model with Uncertainty

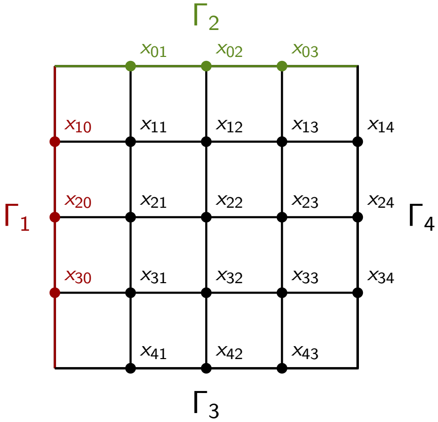
- 2-dimensional heat distribution
  motivated by [Benner/Saak '05]
- boundary control by a cooling fluid with an uncertain spraying intensity

$$
\begin{aligned}
\Omega &= (0,1)\times(0,1) & \\
x_t &= \Delta x & \text{in } \Omega \\
n\cdot\nabla x &= (0.5 + d\omega_1)x & \text{on } \Gamma_1 \\
x &= u & \text{on } \Gamma_2 \\
x &= 0 & \text{on } \Gamma_3, \Gamma_4
\end{aligned}
$$

- spatial discretization $k \times k$-grid

$$\Rightarrow dx \approx Axdt + Nxd\omega_i + Budt$$

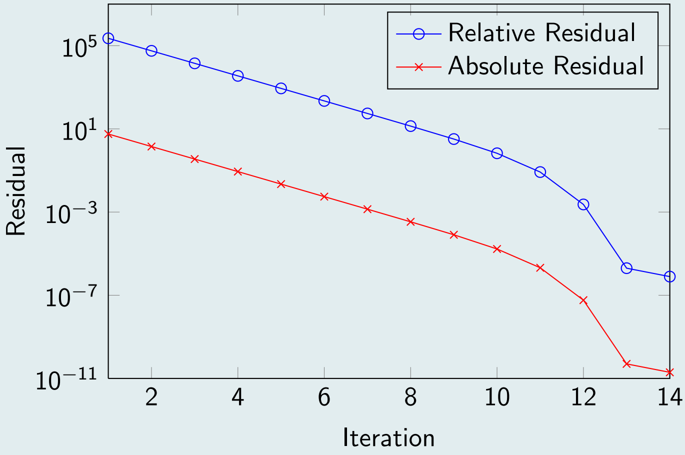
- output: $C = \frac{1}{k^2}\begin{bmatrix}1 & \ldots & 1\end{bmatrix}$



Generalized alternating directions iteration (ADI)

Numerical Example: A Heat Transfer Model with Uncertainty

Conv. history for bilinear low-rank ADI method ($n = 40,000$)



Bilinear Lyapunov Equations

Generalizing the Extended Krylov Subspace Method (EKSM) [Simoncini '07]

Low-rank solutions of the generalized Lyapunov equation now may be obtained by projecting the original equation onto a suitable smaller subspace $\mathcal{V} = \text{span}(V)$, $V \in \mathbb{R}^{n \times k}$, with $V^T V = I$.

In more detail, solve

$$\left(V^T A V\right)\hat{X} + \hat{X}\left(V^T A^T V\right) + \left(V^T N V\right)\hat{X}\left(V^T N^T V\right) + \left(V^T B\right)\left(B^T V\right) = 0$$

and prolongate $X \approx V\hat{X}V^T$.

Bilinear Lyapunov Equations

Generalizing the Extended Krylov Subspace Method (EKSM) [Simoncini '07]

Low-rank solutions of the generalized Lyapunov equation now may be obtained by projecting the original equation onto a suitable smaller subspace $\mathcal{V} = \mathrm{span}(V)$, $V \in \mathbb{R}^{n \times k}$, with $V^T V = I$.

In more detail, solve

$$\left(V^T A V\right) \hat{X} + \hat{X} \left(V^T A^T V\right) + \left(V^T N V\right) \hat{X} \left(V^T N^T V\right) + \left(V^T B\right) \left(B^T V\right) = 0$$

and prolongate $X \approx V \hat{X} V^T$.

For this, one might use the extended Krylov subspace method (EKSM) algorithm in the following way:

Bilinear Lyapunov Equations

Generalizing the Extended Krylov Subspace Method (EKSM) [Simoncini '07]

Low-rank solutions of the generalized Lyapunov equation now may be obtained by projecting the original equation onto a suitable smaller subspace $\mathcal{V} = \text{span}(V)$, $V \in \mathbb{R}^{n \times k}$, with $V^T V = I$.

In more detail, solve

$$\left(V^T A V\right) \hat{X} + \hat{X} \left(V^T A^T V\right) + \left(V^T N V\right) \hat{X} \left(V^T N^T V\right) + \left(V^T B\right) \left(B^T V\right) = 0$$

and prolongate $X \approx V \hat{X} V^T$.

For this, one might use the extended Krylov subspace method (EKSM) algorithm in the following way:

$$V_1 = \begin{bmatrix} B & A^{-1} B \end{bmatrix},$$
$$V_r = \begin{bmatrix} A V_{r-1} & A^{-1} V_{r-1} & N V_{r-1} \end{bmatrix}, \quad r = 2, 3, \ldots$$

Bilinear Lyapunov Equations

Generalizing the Extended Krylov Subspace Method (EKSM) [SIMONCINI '07]

Low-rank solutions of the generalized Lyapunov equation now may be obtained by projecting the original equation onto a suitable smaller subspace $\mathcal{V} = \text{span}(V)$, $V \in \mathbb{R}^{n \times k}$, with $V^T V = I$.

In more detail, solve

$$\left(V^T A V\right) \hat{X} + \hat{X} \left(V^T A^T V\right) + \left(V^T N V\right) \hat{X} \left(V^T N^T V\right) + \left(V^T B\right) \left(B^T V\right) = 0$$

and prolongate $X \approx V \hat{X} V^T$.

For this, one might use the extended Krylov subspace method (EKSM) algorithm in the following way:

$$\begin{aligned} V_1 &= \begin{bmatrix} B & A^{-1}B \end{bmatrix}, \\ V_r &= \begin{bmatrix} AV_{r-1} & A^{-1}V_{r-1} & NV_{r-1} \end{bmatrix}, \quad r = 2, 3, \ldots \end{aligned}$$

However, criteria like dissipativity of A for the linear case which ensure solvability of the projected equation have to be further investigated.

Bilinear EKSM

Residual computation in $\mathcal{O}(k^3)$

Theorem

Let $V_i \in \mathbb{R}^{n\times k_i}$ be the extendend Krylov matrix after i generalized EKSM steps. Denote the residual associated with the approximate solution $X_i = V_i\hat{X}_iV_i^T$ by

$$R_i := AX_i + X_iA^T + NX_iN^T + BB^T,$$

where $\hat{X}_i$ is the solution of the reduced bilinear Lyapunov equation

$$V_i^TAV_i\hat{X}_i + \hat{X}_iV_i^TA^TV_i + V_i^TNV_i\hat{X}_iV_i^TN^TV_i + V_i^TBB^TV_i = 0.$$

Then:

- $\mathrm{range}\,(R_i) \subset \mathrm{range}\,(V_{i+1})$,
- $\|R_i\| = \|V_{i+1}^TR_iV_{i+1}\|$ for the Frobenius and spectral norms.

Bilinear EKSM

Residual computation in $\mathcal{O}(k^3)$

Theorem

Let $V_i \in \mathbb{R}^{n \times k_i}$ be the extendend Krylov matrix after i generalized EKSM steps. Denote the residual associated with the approximate solution $X_i = V_i \hat{X}_i V_i^T$ by

$$R_i := AX_i + X_i A^T + NX_i N^T + BB^T,$$

where $\hat{X}_i$ is the solution of the reduced bilinear Lyapunov equation

$$V_i^T A V_i \hat{X}_i + \hat{X}_i V_i^T A^T V_i + V_i^T N V_i \hat{X}_i V_i^T N^T V_i + V_i^T BB^T V_i = 0.$$

Then:

- $\mathrm{range}\,(R_i) \subset \mathrm{range}\,(V_{i+1})$,
- $\|R_i\| = \|V_{i+1}^T R_i V_{i+1}\|$ for the Frobenius and spectral norms.

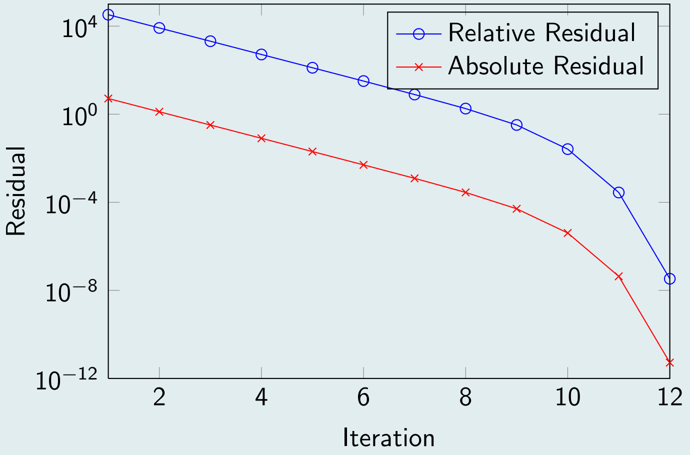
Remarks:

- Residual evaluation only requires quantities needed in $i+1$st projection step plus $\mathcal{O}(k_{i+1}^3)$ operations.
- No Hessenberg structure of reduced system matrix that allows to simplify residual expression as in standard Lyapunov case!

Bilinear EKSM

Numerical Example: A Heat Transfer Model with Uncertainty

Convergence history for bilinear EKSM variant ($n = 6,400$)



Bilinear Lyapunov Equations

Tensorized Krylov subspace methods

Another possibility is to iteratively solve the linear system

$$(I_n \otimes A + A \otimes I_n + N \otimes N)\,\mathrm{vec}(P) = -\,\mathrm{vec}(BB^T),$$

with a fixed number of ADI iteration steps used as a preconditioner $\mathcal{M}$

$$\mathcal{M}^{-1}(I_n \otimes A + A \otimes I_n + A_1 \otimes A_1)\,\mathrm{vec}(P) = -\mathcal{M}^{-1}\,\mathrm{vec}(BB^T).$$

We implemented this approach for PCG and BiCGstab.

Updates like $X_{k+1} \leftarrow X_k + \omega_k P_k$ require truncation operator to preserve low-order structure.

Note, that the low-rank factorization $X \approx ZZ^T$ has to be replaced by $X \approx ZDZ^T$, D possibly indefinite.

Similar to more general tensorized Krylov solvers, see [Kressner/Tobler '10/'12].

Tensorized Krylov subspace methods

Vanilla implementation of tensor-PCG for bilinear matrix equations

Algorithm 1 Preconditioned CG method for $\mathcal{A}(X) - \mathcal{B}$

Input: Matrix functions $\mathcal{A}, \mathcal{M} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, low rank factor B of right-hand side $\mathcal{B} = -BB^T$. Truncation operator $\mathcal{T}$ w.r.t. relative accuracy ϵ_{rel}.

Output: Low rank approximation $X = LDL^T$ with $||\mathcal{A}(X) - \mathcal{B}||_F \leq$ tol.

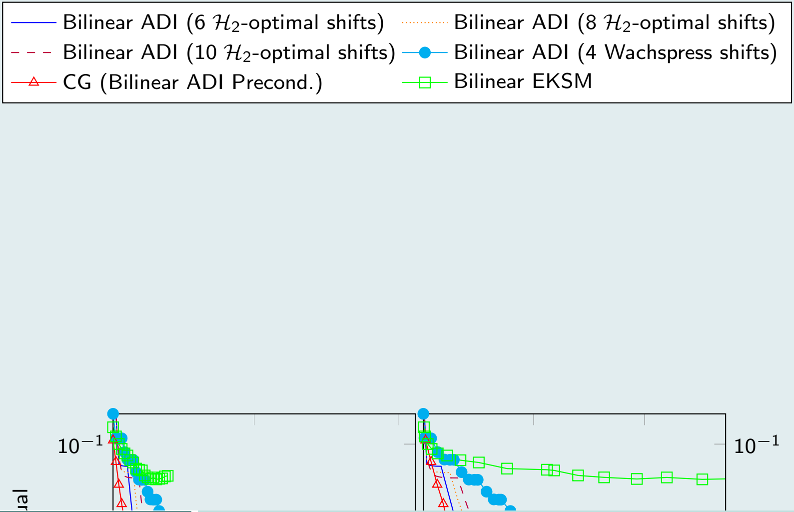
1: $X_0 = 0$, $R_0 = \mathcal{B}$, $Z_0 = \mathcal{M}^{-1}(R_0)$, $P_0 = Z_0$, $Q_0 = \mathcal{A}(P_0)$, $\xi_0 = \langle P_0, Q_0\rangle$, $k = 0$
2: **while** $||R_k||_F >$ tol **do**
3: $\quad \omega_k = \frac{\langle R_k, P_k\rangle}{\xi_k}$
4: $\quad X_{k+1} = X_k + \omega_k P_k$, $\qquad X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$
5: $\quad R_{k+1} = \mathcal{B} - \mathcal{A}(X_{k+1})$, $\qquad$ *Optionally:* $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$
6: $\quad Z_{k+1} = \mathcal{M}^{-1}(R_{k+1})$
7: $\quad \beta_k = -\frac{\langle Z_{k+1}, Q_k\rangle}{\xi_k}$
8: $\quad P_{k+1} = Z_{k+1} + \beta_k P_k$, $\qquad P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$
9: $\quad Q_{k+1} = \mathcal{A}(P_{k+1})$, $\qquad$ *Optionally:* $Q_{k+1} \leftarrow \mathcal{T}(Q_{k+1})$
10: $\quad \xi_{k+1} = \langle P_{k+1}, Q_{k+1}\rangle$
11: $\quad k = k + 1$
12: **end while**
13: $X = X_k$

Here, $\mathcal{A} : X \to AX + XA^T + NXN^T$, $\mathcal{M}$: ℓ steps of (bilinear) ADI, both in low-rank ("ZDZ^T" format).

Comparison of methods

Heat equation with boundary control

Comparison of low rank solution methods for $n = 562,500$.

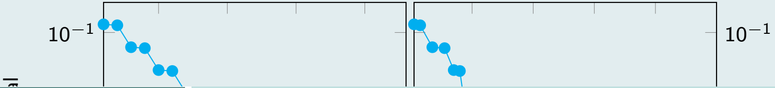


Comparison of methods

Fokker-Planck equation

Comparison of low rank solution methods for $n = 10,000$.

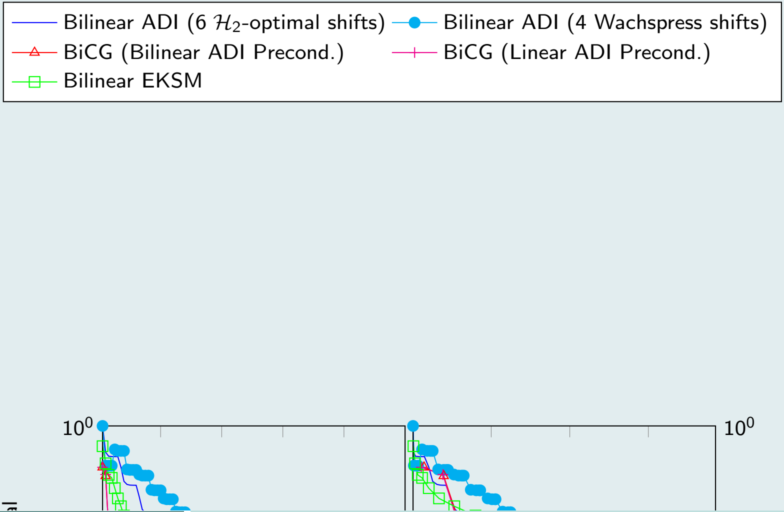
Bilinear ADI (2 $\mathcal{H}_2$-optimal shifts)	Bilinear ADI (2 Wachspress shifts)
BiCG (Bilinear ADI Precond.)	BiCG (Linear ADI Precond.)
Bilinear EKSM	



Comparison of methods

RC circuit simulation

Comparison of low rank solution methods for $n = 250,000$.



Comparison of methods

Comparison of CPU times

	Heat equation	RC circuit	Fokker-Planck
Bilin. ADI 2 $\mathcal{H}_2$ shifts	-	-	1.733 (1.578)
Bilin. ADI 6 $\mathcal{H}_2$ shifts	144,065 (2,274)	20,900 (3091)	-
Bilin. ADI 8 $\mathcal{H}_2$ shifts	135,711 (3,177)	-	-
Bilin. ADI 10 $\mathcal{H}_2$ shifts	33,051 (4,652)	-	-
Bilin. ADI 2 Wachspress shifts	-	-	6.617 (4.562)
Bilin. ADI 4 Wachspress shifts	41,883 (2,500)	18,046 (308)	-
CG (Bilin. ADI precond.)	15,640	-	-
BiCG (Bilin. ADI precond.)	-	16,131	11.581
BiCG (Linear ADI precond.)	-	12,652	9.680
EKSM	7,093	19,778	8.555

Numbers in brackets: computation of shift parameters.

Application to Parametric MOR

Fast simulation of cyclic voltammogramms [Feng/Koziol/Rudnyi/Korvink '06]

$$E\dot{x}(t) = (A + p_1(t)A_1 + p_2(t)A_2)x(t) + B,$$
$$y(t) = Cx(t), \quad x(0) = x_0 \neq 0,$$

- Rewrite as system with zero initial condition,
- FE model: $n = 16,912, \ m = 3, \ q = 1,$
- $p_j \in [0, 10^9]$ time-varying voltage functions,
- transfer function $H(i\omega, p_1, p_2)$,
- reduced system dimension $r = 67$,
- $\max\limits_{\substack{\omega \in \{\omega_{min},\dots,\omega_{max}\} \\ p_j \in \{p_{min},\dots,p_{max}\}}} \frac{||H - \hat{H}||_2}{||H||_2} < 6 \cdot 10^{-4}$,
- evaluation times: FOM 4.5h, ROM 38s $\rightsquigarrow$ speed-up factor ≈ 426.

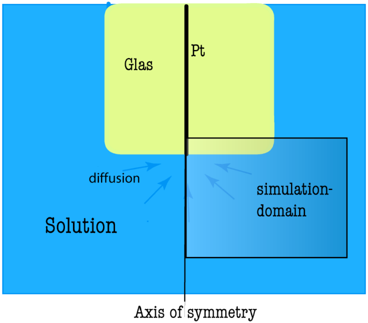
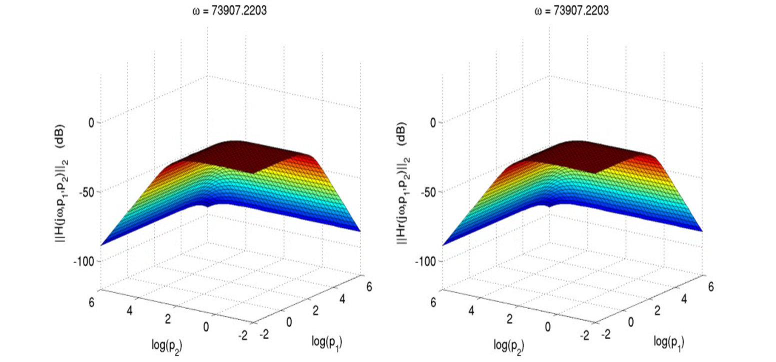


Figure: [Feng et al. '06]

Application to Parametric MOR

Fast simulation of cyclic voltammogramms [Feng/Koziol/Rudnyi/Korvink '06]

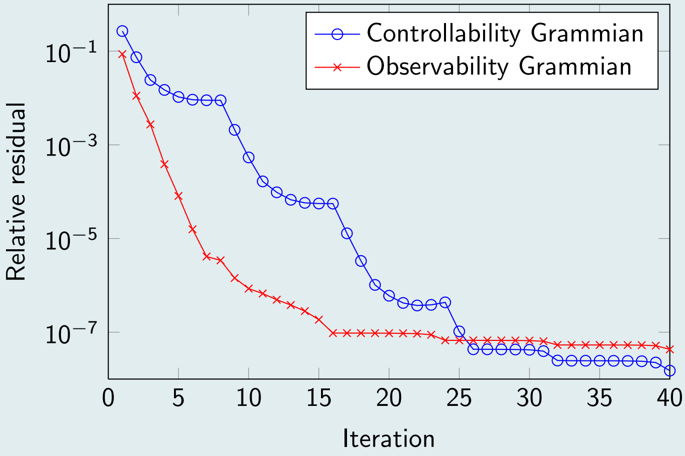
Original. . . **and reduced-order model.**



Application to Parametric MOR

Fast simulation of cyclic voltammogramms [Feng/Koziol/Rudnyi/Korvink '06]

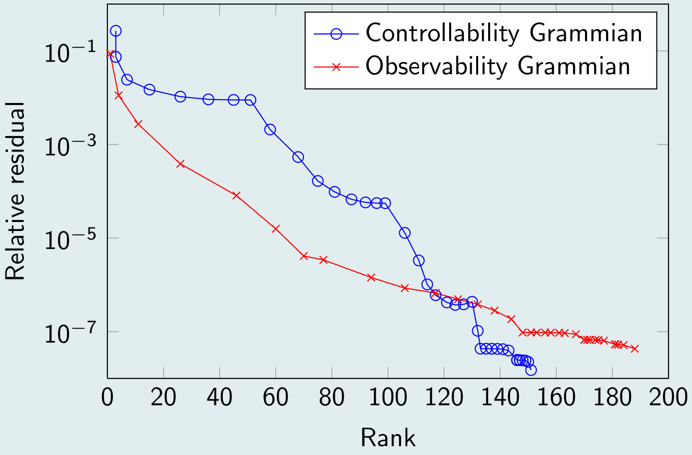
Convergence history for bilinear ADI iteration



Application to Parametric MOR

Fast simulation of cyclic voltammogramms [Feng/Koziol/Rudnyi/Korvink '06]

Rank increase for bilinear ADI iteration



Application to Parametric MOR

2D model of an anemometer [Baur et al. '10]

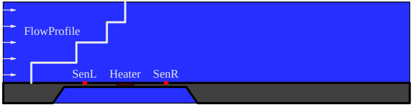


Figure: [Baur et al. '10]

Consider an anemometer, a flow sensing device located on a membrane used in context of minimizing heat dissipation.

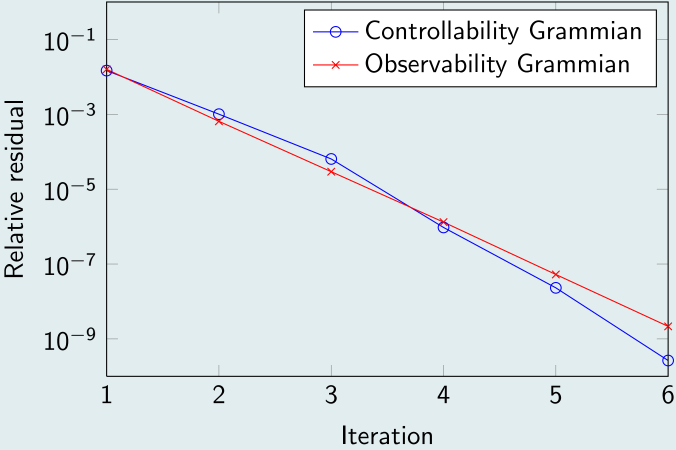
$$E\dot{x}(t) = (A + pA_1)x(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0,$$

- FE model: $n = 29,008$, $m = 1$, $q = 3$,
- $p_1 \in [0, 1]$ fluid velocity,
- transfer function $H(i\omega, p_1)$, reduced system dimension $r = 146$,
- $\max\limits_{\substack{\omega \in \{\omega_{min},\dots,\omega_{max}\} \\ p_1 \in \{p_{min},\dots,p_{max}\}}} \frac{\|H(\omega,p) - \hat{H}(\omega,p)\|_2}{\|H(\omega,p)\|_2} < 3 \cdot 10^{-5}$,
- evaluation times: FOM 51min, ROM 21s.

Application to Parametric MOR

2D model of an anemometer [Baur et al. '10]

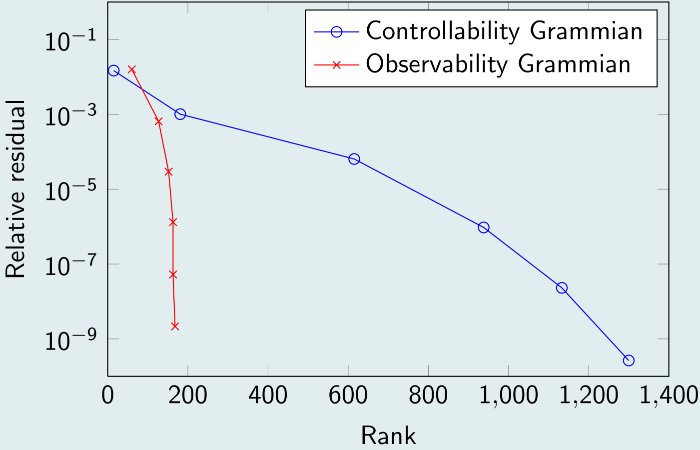
Convergence history for preconditioned BiCGstab



Application to Parametric MOR

2D model of an anemometer [Baur et al. '10]

Rank increase for preconditioned BiCGstab



Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma : \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

Here, $A, N \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{q\times n}$ and $d\omega(t)$ are white noise processes associated with a Wiener process $\omega(t)$.

The SDE formalism is merely a notation for

$$x(t) = x(0) + \int_0^t Ax(\tau)d\tau + \int_0^t Nx(\tau)d\omega + \int_0^t Bu(\tau)d\tau,$$

with $d\omega_i$ denoting the Itô integral.

Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma: \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

Balanced truncation for linear SDEs requires the same steps as for deterministic systems, instead of standard Lyapunov equations need to solve again

$$AP + PA^T + NPN^T + BB^T = 0,$$
$$A^TQ + QA^T + N^TQN + C^TC = 0.$$

Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma : \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

Balanced truncation for linear SDEs requires the same steps as for deterministic systems, instead of standard Lyapunov equations need to solve again

$$AP + PA^T + NPN^T + BB^T = 0,$$
$$A^T Q + QA^T + N^T QN + C^T C = 0.$$

Can show stability preservation under certain assumptions, but no error bound!

Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma : \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

Balanced truncation for linear SDEs requires the same steps as for deterministic systems, instead of standard Lyapunov equations need to solve again

$$AP + PA^T + NPN^T + BB^T = 0,$$
$$A^T Q + QA^T + N^T QN + C^T C = 0.$$

Can show stability preservation under certain assumptions, but no error bound!
Alternative: Applying balanced truncation using $\tilde{P}$ and Q, where

$$A\tilde{P} + \tilde{P}A^T + \tilde{P}N\tilde{P}^{-1}N^T\tilde{P} + BB^T = 0,$$

we obtain desired error bound $\|y - \hat{y}\|_{L^2_\omega} \leq 2\left(\sum_{j=r+1}^{n} \sigma_j\right) \|u\|_{L^2_\omega}$ [B./Damm '12].

Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma : \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

Balanced truncation for linear SDEs requires the same steps as for deterministic systems, instead of standard Lyapunov equations need to solve again

$$AP + PA^T + NPN^T + BB^T = 0,$$
$$A^T Q + QA^T + N^T QN + C^T C = 0.$$

Can show stability preservation under certain assumptions, but no error bound!
Alternative: Applying balanced truncation using $\tilde{P}$ and Q, where

$$A\tilde{P} + \tilde{P}A^T + \tilde{P}N\tilde{P}^{-1}N^T\tilde{P} + BB^T = 0,$$

we obtain desired error bound $\|y-\hat{y}\|_{L^2_\omega} \leq 2\left(\sum_{j=r+1}^{n} \sigma_j\right)\|u\|_{L^2_\omega}$ [B./Damm '12].
Problem: no satisfactory solution method for new nonlinear matrix equation!
Note: using $\hat{P} := \tilde{P}^{-1}$, we obtain an algebraic Bernoulli equation:

$$\hat{P}A + A^T\hat{P} + N\hat{P}N^T + \hat{P}BB^T\hat{P} = 0.$$

Stochastic Systems

Itô-type linear stochastic differential equations (SDE):

$$\Sigma : \quad \begin{cases} dx(t) = Ax(t)dt + Nx(t)d\omega(t) + Bu(t)dt, \\ y(t) = Cx(t), \quad x(0) = x_0. \end{cases}$$

In analogy to LGQ BT for deterministic systems, could use solutions of stochastic LQG equations [Wonham '68]:

$$\begin{aligned} 0 &= AP + PA^T + NQN^T - PC^TCP + BB^T, \\ 0 &= A^TQ + QA + N^TQN - QBB^TQ + C^TC. \end{aligned}$$

- Solution for large-scale problems using Newton-ADI or Newton-EKSM [B./Breiten ILAS 2011].
- No results regarding properties of reduced-order model.
- Might also use these equations for "LQG BT" for bilinear systems, but: these are not LQG equations from LQG design for bilinear systems; there, P, Q are state-dependent as, e.g., $B \to B + Nx$!

Conclusions and Outlook

- Model reduction for bilinear and stochastic systems leads to the solution of generalized ("bilinear") Lyapunov equations.
- Special versions of balanced truncation for stochastic systems lead to nonlinear matrix equations.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- Under certain assumptions, we can expect the existence of low-rank approximations to the solution of generalized Lyapunov equations.
- Solutions strategies via extending the ADI iteration to bilinear systems and EKSM as well as using preconditioned iterative solvers like CG or BiCGstab up to dimensions $n \sim 500,000$ in MATLAB®.
- Optimal choice of shift parameters for ADI is a nontrivial task.
- What about the singular value decay in case of N being full rank?

References

P. Benner and T. Breiten.
On $\mathcal{H}_2$ model reduction of linear parameter-varying systems.
Proceedings in Applied Mathematics and Mechanics 11:805-806, 2011.

P. Benner and T. Breiten.
On optimality of interpolation-based low-rank approximations of large-scale matrix equations.
Preprint MPIMD/11-10, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, December 2011.

P. Benner and T. Breiten.
Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems.
SIAM Journal on Matrix Analysis and Applications 33(3):859–885, 2012.

P. Benner and T. Breiten.
Low rank methods for a class of generalized Lyapunov equations and related issues.
Numerische Mathematik, to appear.

P. Benner and T. Damm
Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems.
SIAM Journal on Control and Optimization 49(2):686–711, 2011.

T. Damm.
Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations.
Numerical Linear Algebra with Applications 15(9):853–871, 2008.

L. Grasedyck.
Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure.
Computing 72(3–4):247–265, 2004.

(Upcoming) preprints available at

http://www.mpi-magdeburg.mpg.de/preprints/index.php