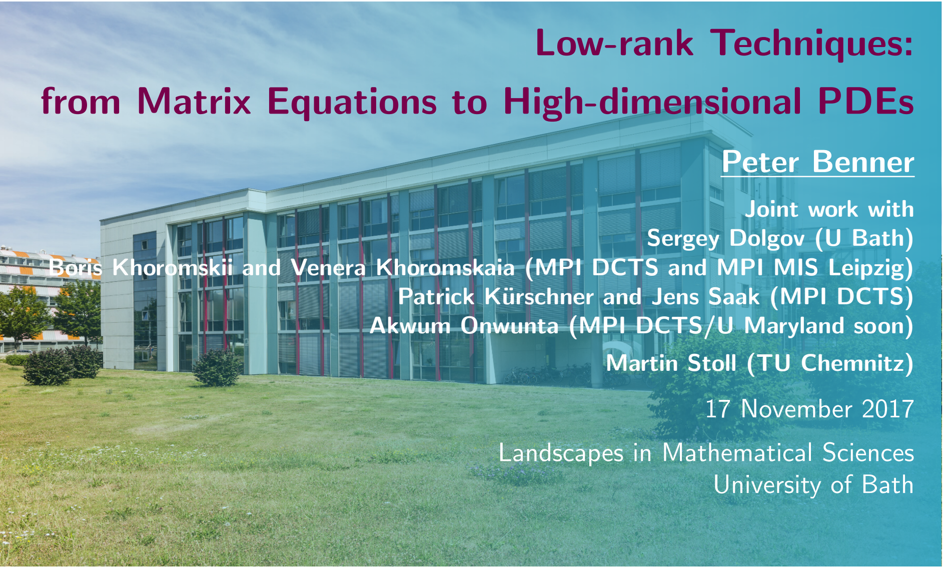
MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC COMPUTATIONAL METHODS IN SYSTEMS AND CONTROL THEORY



Low-rank Techniques: from Matrix Equations to High-dimensional PDEs

Peter Benner

Joint work with
Sergey Dolgov (U Bath)
Boris Khoromskii and Venera Khoromskaia (MPI DCTS and MPI MIS Leipzig)
Patrick Kürschner and Jens Saak (MPI DCTS)
Akwum Onwunta (MPI DCTS/U Maryland soon)
Martin Stoll (TU Chemnitz)

17 November 2017

Landscapes in Mathematical Sciences
University of Bath

The Max Planck Institute (MPI) in Magdeburg

The Max Planck **Society**

- operates 84 institutes — 79 in Germany, 2 in Italy, 1 each in The Netherlands, Luxembourg, and the USA;
- has $\sim$ 23,000 employees;
- had 18 Nobel Laureates since 1948.

"The first MPI in engineering. . . "

MPI Magdeburg

- founded 1998
- 4 departments (directors)
- 9 research groups
- budget $\sim$ 15 Mio. EUR
- $\sim$ 230 employees
- 130 scientists,
- doing research in
  - biotechnology
  - chemical engineering
  - process engineering
  - energy conversion
  - applied math

Outline

1. Motivation
2. Solving Large-Scale Sylvester and Lyapunov Equations
   - Some Basics
   - The Low-Rank Structure
   - LR-ADI Method
   - The New LR-ADI Applied to Lyapunov Equations
3. From Matrix Equations to PDEs in d Dimensions
   - The Curse of Dimensionality
   - Tensor Techniques
   - Numerical Examples

Outline

1. Motivation
2. Solving Large-Scale Sylvester and Lyapunov Equations
3. From Matrix Equations to PDEs in d Dimensions

Model Reduction in Frequency Domain

Approximate the linear control system

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m}, \\ y &= Cx + Du, \qquad C \in \mathbb{R}^{q\times n},\ D \in \mathbb{R}^{q\times m}, \end{aligned}$$

by reduced-order system

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

of order $r \ll n$, such that

$$\| y - \hat{y} \| = \| Gu - \hat{G}u \| \leq \| G - \hat{G} \| \cdot \| u \| < \text{tolerance} \cdot \| u \|,$$

where

$$G(s) = C(sI_n - A)^{-1}B + D, \quad \hat{G}(s) = \hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}$$

are the associated transfer functions of the system.

Balanced Truncation [Moore '81]

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

  is balanced, if system Gramians = solutions P, Q of 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 [Moore '81]

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

  is balanced, if system Gramians = solutions P, Q of 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.

Balanced Truncation [Moore '81]

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

  is balanced, if system Gramians = solutions P, Q of 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.
- Compute balanced realization of Σ 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 [Moore '81]

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

  is balanced, if system Gramians = solutions P, Q of 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 Σ 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 $\rightsquigarrow (\hat{A}, \hat{B}, \hat{C}) = (A_{11}, B_1, C_1)$.

Prelude — 20 years of working in low-rank methods

Balanced Truncation **[Moore '81]**

. . . is one of the greatest model reduction techniques for linear systems

Balanced Truncation [Moore '81]

...is one of the greatest model reduction techniques for linear systems since

- the reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$;

Balanced Truncation [Moore '81]

. . . is one of the greatest model reduction techniques for linear systems since

- the reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$;
- it allows adaptive choice of r via computable error bound (there is a free lunch!):

$$\| y - \hat{y} \|_2 \leq \| G - \hat{G} \|_{\mathcal{H}_\infty} \| u \|_2 \leq \left(2 \sum\nolimits_{k=r+1}^{n} \sigma_k \right) \| u \|_2 .$$

Balanced Truncation [Moore '81]

...is one of the greatest model reduction techniques for linear systems!

But: as suggested in [Moore '81], it is an $\mathcal{O}(n^3)$ **method!**
(Bottleneck: solution of the Lyapunov equations, followed by SVD of their Cholesky factors to determine necessary parts of T and T^{-1}.)

Balanced Truncation [Moore '81]

... is one of the greatest model reduction techniques for linear systems!

But: as suggested in [Moore '81], it is an $\mathcal{O}(n^3)$ **method!**
(Bottleneck: solution of the Lyapunov equations, followed by SVD of their Cholesky factors to determine necessary parts of T and T^{-1}.)

In **1997**, we derived a method to compute **low-rank factors** $S, R \in \mathbb{R}^{n\times\ell}$, $\ell \ll n$, such that $P \approx SS^T$, $Q \approx RR^T$ [1] and then showed that the truncation could be obtained via cheap $\ell_S \times \ell_R$ SVD of $S^T R$ [2]!

(Some similar ideas: low-rank Lyapunov solver [Saad '90, Jaimoukha/Kasenally '94], small-scale SVD in BT [Penzl '98, Li/White '99].)

P. Benner and E.S. Quintana-Ortí. Solving stable generalized Lyapunov equations with the matrix sign function. Numerical Algorithms, 20(1):75–100, 1999. (Preprint SFB393 97–23)

P. Benner, E.S. Quintana-Ortí, and G. Quintana-Ortí. Balanced truncation model reduction of large-scale dense systems on parallel computers. Mathematical and Computer Modeling of Dynamical Systems, 6(4):383–405, 2000.

Outline

Sylvester equation

James Joseph Sylvester
(September 3, 1814 – March 15, 1897)

$$AX + XB = C.$$

Sylvester equation

James Joseph Sylvester
(September 3, 1814 – March 15, 1897)

$$AX + XB = C.$$

Lyapunov equation

Alexander Michailowitsch Ljapunow
(June 6, 1857 – November 3, 1918)

$$AX + XA^T = C, \quad C = C^T.$$

Generalizations of Sylvester $(AX + XB = C)$ and Lyapunov $(AX + XA^T = C)$ Equations

Generalized Sylvester equation:

$$AXD + EXB = C.$$

Generalized Lyapunov equation:

$$AXE^T + EXA^T = C, \quad C = C^T.$$

Stein equation:

$$X - AXB = C.$$

(Generalized) discrete Lyapunov/Stein equation:

$$EXE^T - AXA^T = C, \quad C = C^T.$$

Generalizations of Sylvester ($AX + XB = C$) and Lyapunov ($AX + XA^T = C$) Equations

Generalized Sylvester equation:

$$AXD + EXB = C.$$

Generalized Lyapunov equation:

$$AXE^T + EXA^T = C, \quad C = C^T.$$

Stein equation:

$$X - AXB = C.$$

(Generalized) discrete Lyapunov/Stein equation:

$$EXE^T - AXA^T = C, \quad C = C^T.$$

Note:

- Consider only regular cases, having a unique solution!
- Solutions of symmetric cases are symmetric, $X = X^T \in \mathbb{R}^{n \times n}$; otherwise, $X \in \mathbb{R}^{n \times \ell}$ with $n \neq \ell$ in general.

Applications

- Gramian-based model order reduction
- Linear stability analysis
- Continuation algorithms for detecting Hopf bifurcations
- Determining metastable equilibrium points of stochastic differential equations
- Block-triangularization of matrices and matrix pencils
- Computing covariance matrices, e.g., in biochemical reaction networks
- Numerical solution of fractional partial differential equations
- Solving algebraic Riccati equations via Newton's method
- . . .

Exemplarily, consider the generalized Sylvester equation

$$AXD + EXB = C. \tag{1}$$

Existence of Solutions of Linear Matrix Equations

Exemplarily, consider the generalized Sylvester equation

$$AXD + EXB = C. \tag{1}$$

Vectorization (using Kronecker product) $\rightsquigarrow$ representation as linear system:

$$\big(\underbrace{D^T \otimes A + B^T \otimes E}_{=:\mathcal{A}}\big)\underbrace{\mathrm{vec}(X)}_{=:x} = \underbrace{\mathrm{vec}(C)}_{=:c} \qquad \Longleftrightarrow \qquad \mathcal{A}x = c.$$

Existence of Solutions of Linear Matrix Equations

Exemplarily, consider the generalized Sylvester equation

$$AXD + EXB = C. \tag{1}$$

Vectorization (using Kronecker product) $\rightsquigarrow$ representation as linear system:

$$\big(\underbrace{D^T \otimes A + B^T \otimes E}_{=:\mathcal{A}}\big)\underbrace{\mathrm{vec}(X)}_{=:x} = \underbrace{\mathrm{vec}(C)}_{=:c} \qquad \Longleftrightarrow \qquad \mathcal{A}x = c.$$

$\Longrightarrow$ "(1) has a unique solution $\Longleftrightarrow$ $\mathcal{A}$ is nonsingular"

Existence of Solutions of Linear Matrix Equations

Exemplarily, consider the generalized Sylvester equation

$$AXD + EXB = C. \tag{1}$$

Vectorization (using Kronecker product) $\rightsquigarrow$ representation as linear system:

$$\underbrace{\left(D^T \otimes A + B^T \otimes E\right)}_{=:\mathcal{A}} \underbrace{\mathrm{vec}(X)}_{=:x} = \underbrace{\mathrm{vec}(C)}_{=:c} \qquad \Longleftrightarrow \qquad \mathcal{A}x = c.$$

$\Longrightarrow$ "(1) has a unique solution $\Longleftrightarrow$ $\mathcal{A}$ is nonsingular"

Lemma

$$\Lambda(\mathcal{A}) = \{\alpha_j + \beta_k \mid \alpha_j \in \Lambda(A, E),\ \beta_k \in \Lambda(B, D)\}.$$

Hence, (1) has unique solution $\Longleftrightarrow \Lambda(A, E) \cap -\Lambda(B, D) = \emptyset$.

Exemplarily, consider the generalized Sylvester equation

$$AXD + EXB = C. \tag{1}$$

Vectorization (using Kronecker product) $\rightsquigarrow$ representation as linear system:

$$\big(\underbrace{D^T \otimes A + B^T \otimes E}_{=:\mathcal{A}}\big)\underbrace{\operatorname{vec}(X)}_{=:x} = \underbrace{\operatorname{vec}(C)}_{=:c} \qquad \Longleftrightarrow \qquad \mathcal{A}x = c.$$

$\Longrightarrow$ "(1) has a unique solution $\Longleftrightarrow$ $\mathcal{A}$ is nonsingular"

Lemma

$$\Lambda(\mathcal{A}) = \{\alpha_j + \beta_k \mid \alpha_j \in \Lambda(A, E),\ \beta_k \in \Lambda(B, D)\}.$$

Hence, (1) has unique solution $\Longleftrightarrow \Lambda(A, E) \cap -\Lambda(B, D) = \emptyset$.

Example: Lyapunov equation $AX + XA^T = C$ has unique solution

$$\Longleftrightarrow \nexists\, \mu \in \mathbb{C}\ :\ \pm\mu \in \Lambda(A).$$

The Low-Rank Structure

Sylvester Equations

Find $X \in \mathbb{R}^{n\times m}$ solving

$$AX - XB = FG^T,$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, $F \in \mathbb{R}^{n\times r}$, $G \in \mathbb{R}^{m\times r}$.

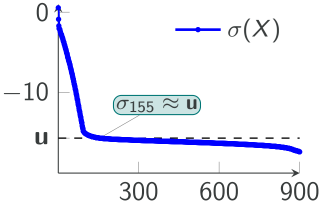
If n, m large, but $r \ll n, m$

$\rightsquigarrow$ X has a small numerical rank.

[Penzl 1999, Grasedyck 2004, Antoulas/Sorensen/Zhou 2002]

$$\operatorname{rank}(X, \tau) = f \ll \min(n, m)$$

singular values of 1600×900 example



$\rightsquigarrow$ Compute low-rank solution factors $Z \in \mathbb{R}^{n\times f}$, $Y \in \mathbb{R}^{m\times f}$, $D \in \mathbb{R}^{f\times f}$, such that $X \approx ZDY^T$ with $f \ll \min(n, m)$.

The Low-Rank Structure

Lyapunov Equations

Find $X \in \mathbb{R}^{n \times n}$ solving

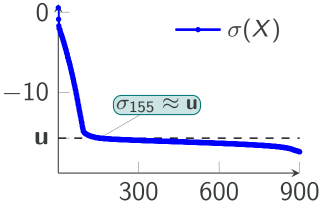
$$AX + XA^T = -FF^T,$$

where $A \in \mathbb{R}^{n \times n}$, $F \in \mathbb{R}^{n \times r}$.

If n large, but $r \ll n$
$\rightsquigarrow$ X has a small numerical rank.
[Penzl 1999, Grasedyck 2004, Antoulas/Sorensen/Zhou 2002]

$$\operatorname{rank}(X, \tau) = f \ll n$$

singular values of 1600×900 example



$\rightsquigarrow$ Compute low-rank solution factors $Z \in \mathbb{R}^{n \times f}$, $D \in \mathbb{R}^{f \times f}$, such that $X \approx ZDZ^T$ with $f \ll n$.

Theoretical Foundation of the Low-rank Structure

Lyapunov case: [Grasedyck '04]

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

Theoretical Foundation of the Low-rank Structure

Lyapunov case: [Grasedyck '04]

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

For stable M, i.e., $\Lambda(M) \subset \mathbb{C}^-$, 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

Theoretical Foundation of the Low-rank Structure

Lyapunov case: [Grasedyck '04]

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

For stable M, i.e., $\Lambda(M) \subset \mathbb{C}^-$, 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

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

Theoretical Foundation of the Low-rank Structure

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

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

Theoretical Foundation of the Low-rank Structure

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,

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

Theoretical Foundation of the Low-rank Structure

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,

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

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

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

Extended to $AX + XA^T + \sum_{j=1}^m N_j X N_j^T + BB^T = 0.$ [B./Breiten 2012]

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Definition

The matrix sign function of $M \in \mathbb{R}^{n \times n}$ with no purely imaginary eigenvalues is

$$\operatorname{sign}(M) = \operatorname{sign}\left(T \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} T^{-1} \right) = T \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} T^{-1}$$

with $J_\pm$ containing all Jordan blocks of M corresponding to eigenvalues with positive/negative real parts.

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Definition

The matrix sign function of $M \in \mathbb{R}^{n \times n}$ with no purely imaginary eigenvalues is

$$\operatorname{sign}(M) = \operatorname{sign}\left(T \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} T^{-1} \right) = T \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} T^{-1}$$

with $J_\pm$ containing all Jordan blocks of M corresponding to eigenvalues with positive/negative real parts.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix} \right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}$.
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}$.
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Sign function iteration for solving Lyapunov equations

$M_0 = \begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}$, and inversion formula for block-triangular matrices:

$$\begin{aligned} A_{k+1} &\leftarrow \frac{1}{2}(A_k + A_k^{-1}) \\ B_{k+1}B_{k+1} &\leftarrow \frac{1}{2}(B_k B_k^T + A_k^{-1} B_k B_k^T A_k^{-T}) \end{aligned}$$

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}$.
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Sign function iteration for solving Lyapunov equations

$M_0 = \begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}$, and inversion formula for block-triangular matrices:

$$\begin{aligned} A_{k+1} &\leftarrow \frac{1}{2}(A_k + A_k^{-1}) \\ B_{k+1}B_{k+1}^T &\leftarrow \frac{1}{2}\left(B_k B_k^T + A_k^{-1} B_k (A_k^{-1} B_k)^T\right) \end{aligned}$$

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}$.
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Sign function iteration for solving Lyapunov equations

$M_0 = \begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}$, and inversion formula for block-triangular matrices:

$$\begin{aligned} A_{k+1} &\leftarrow \frac{1}{2}(A_k + A_k^{-1}) \\ B_{k+1}B_{k+1}^T &\leftarrow \frac{1}{2}\left(B_k B_k^T + A_k^{-1}B_k(A_k^{-1}B_k)^T\right) \end{aligned}$$

so that $B_k \to \frac{1}{\sqrt{2}}Z$ with $X = ZZ^T$.

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}.$
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Factored sign function iteration for Lyapunov equations [B./Quintana-Ortí 1997/99]

$$\begin{aligned} A_{k+1} &\leftarrow \frac{1}{2}(A_k + A_k^{-1}) \\ B_{k+1} &\leftarrow \frac{1}{\sqrt{2}}[\,B_k,\ A_k^{-1}B_k\,] \end{aligned}$$

Problem: number of columns in B_k doubles each iteration!

Exploiting the Low-Rank Structure I

For simplicity, consider again the Lyapunov equation $AX + XA^T + BB^T = 0$.

Observations

1. $\operatorname{sign}\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) = \begin{bmatrix} -I & 2X \\ 0 & I \end{bmatrix}$.
2. $\operatorname{sign}(M) = \lim_{k\to\infty} M_k$ with $M_{k+1} = \frac{1}{2}(M_k + M_k^{-1})$ if $M_0 = M$.

Factored sign function iteration for Lyapunov equations [B./Quintana-Ortí 1997/99]

$$\begin{aligned} A_{k+1} &\leftarrow \frac{1}{2}(A_k + A_k^{-1}) \\ B_{k+1} &\leftarrow \frac{1}{\sqrt{2}}[\,B_k,\ A_k^{-1}B_k\,] \end{aligned}$$

Problem: number of columns in B_k doubles each iteration!

Cure: **truncation operator** $B_{k+1} \leftarrow \mathcal{T}_\varepsilon\left(\frac{1}{\sqrt{2}}[\,B_k,\ A_k^{-1}B_k\,]\right)$ with, e.g., $\mathcal{T}_\varepsilon$ returning the scaled left singular vectors of the truncated SVD w.r.t. the numerical rank tolerance ε.

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \mathrm{vec}(X) = \mathrm{vec}(FG^T).$$

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \operatorname{vec}(X) = \operatorname{vec}(FG^T).$$

This cannot be used for numerical solutions unless $nm \leq 1,000$ (or so), as

- it requires $\mathcal{O}(n^2m^2)$ of storage;

Exploiting the Low-Rank Structure II

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \operatorname{vec}(X) = \operatorname{vec}(FG^T).$$

This cannot be used for numerical solutions unless $nm \leq 1,000$ (or so), as

- it requires $\mathcal{O}(n^2 m^2)$ of storage;
- direct solver needs $\mathcal{O}(n^3 m^3)$ flops;

Exploiting the Low-Rank Structure II

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \operatorname{vec}(X) = \operatorname{vec}(FG^T).$$

This cannot be used for numerical solutions unless $nm \leq 1,000$ (or so), as

- it requires $\mathcal{O}(n^2m^2)$ of storage;
- direct solver needs $\mathcal{O}(n^3m^3)$ flops;
- low (tensor-)rank of right-hand side is ignored;

Exploiting the Low-Rank Structure II

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \operatorname{vec}(X) = \operatorname{vec}(FG^T).$$

This cannot be used for numerical solutions unless $nm \leq 1,000$ (or so), as

- it requires $\mathcal{O}(n^2 m^2)$ of storage;
- direct solver needs $\mathcal{O}(n^3 m^3)$ flops;
- low (tensor-)rank of right-hand side is ignored;
- in Lyapunov case, symmetry and possible definiteness are not respected.

Exploiting the Low-Rank Structure II

Exploiting the Low-rank Structure for Large and **Sparse** Coefficients

Sylvester equation $AX - XB = FG^T$ is equivalent to linear system of equations

$$\left(I_m \otimes A - B^T \otimes I_n\right) \operatorname{vec}(X) = \operatorname{vec}(FG^T).$$

This cannot be used for numerical solutions unless $nm \leq 1,000$ (or so), as

- it requires $\mathcal{O}(n^2 m^2)$ of storage;
- direct solver needs $\mathcal{O}(n^3 m^3)$ flops;
- low (tensor-)rank of right-hand side is ignored;
- in Lyapunov case, symmetry and possible definiteness are not respected.

Possible solvers:

- Standard Krylov subspace solvers in operator from [Hochbruck, Starke, Reichel, ...]
- Block-Tensor-Krylov subspace methods with truncation [Kressner/Tobler, Bollhöfer/Eppler, B./Breiten, ...]
- Galerkin-type methods based on (extended, rational) Krylov subspace methods [Jaimoukha, Kasenally, Jbilou, Simoncini, Druskin, Knizhermann,...]
- Doubling-type methods [Smith, Chu et al., B./Sadkane/El Khoury, ...]
- ADI methods [Wachspress, Reichel, Li, Penzl, B., Saak, Kürschner, Opmeer/Reis, ...]

LR-ADI Method

Sylvester and Stein equations

Let $\alpha \neq \beta$ with $\alpha \notin \Lambda(B)$, $\beta \notin \Lambda(A)$, then

$$\underbrace{AX - XB = FG^T}_{\text{Sylvester equation}} \quad \Leftrightarrow \quad \underbrace{X = \mathcal{A} X \mathcal{B} + (\beta - \alpha)\mathcal{F}\mathcal{G}^H}_{\text{Stein equation}}$$

with the Cayley-like (Möbius) transformations

$$\mathcal{A} := (A - \beta I_n)^{-1}(A - \alpha I_n), \qquad \mathcal{B} := (B - \alpha I_m)^{-1}(B - \beta I_m),$$
$$\mathcal{F} := (A - \beta I_n)^{-1}F, \qquad \mathcal{G} := (B - \alpha I_m)^{-H}G.$$

$\rightsquigarrow$ fix point iteration

$$X_k = \mathcal{A} X_{k-1} \mathcal{B} + (\beta - \alpha)\mathcal{F}\mathcal{G}^H$$

for $k \geq 1$, $X_0 \in \mathbb{R}^{n \times m}$.

LR-ADI Method

Sylvester and Stein equations

Let $\alpha_\mathbf{k} \neq \beta_\mathbf{k}$ with $\alpha_\mathbf{k} \notin \Lambda(B)$, $\beta_\mathbf{k} \notin \Lambda(A)$, then

$$\underbrace{AX - XB = FG^T}_{\text{Sylvester equation}} \quad \Leftrightarrow \quad \underbrace{X = \mathcal{A}_\mathbf{k} X \mathcal{B}_\mathbf{k} + (\beta_\mathbf{k} - \alpha_\mathbf{k}) \mathcal{F}_\mathbf{k} {\mathcal{G}_\mathbf{k}}^H}_{\text{Stein equation}}$$

with the Cayley-like (Möbius) transformations

$$\mathcal{A}_\mathbf{k} := (A - \beta_\mathbf{k} I_n)^{-1}(A - \alpha_\mathbf{k} I_n), \qquad \mathcal{B}_\mathbf{k} := (B - \alpha_\mathbf{k} I_m)^{-1}(B - \beta_\mathbf{k} I_m),$$

$$\mathcal{F}_\mathbf{k} := (A - \beta_\mathbf{k} I_n)^{-1} F, \qquad \mathcal{G}_\mathbf{k} := (B - \alpha_\mathbf{k} I_m)^{-H} G.$$

$\rightsquigarrow$ **alternating directions implicit (ADI)** iteration

$$X_k = \mathcal{A}_\mathbf{k} X_{k-1} \mathcal{B}_\mathbf{k} + (\beta_\mathbf{k} - \alpha_\mathbf{k}) \mathcal{F}_\mathbf{k} {\mathcal{G}_\mathbf{k}}^H$$

for $k \geq 1$, $X_0 \in \mathbb{R}^{n \times m}$. [Wachspress 1988]

LR-ADI Method

Sylvester ADI iteration [Wachspress 1988]

$$\begin{aligned}
X_k &= \mathcal{A}_k X_{k-1} \mathcal{B}_k + (\beta_k - \alpha_k)\mathcal{F}_k \mathcal{G}_k^H, \\
\mathcal{A}_k &:= (A - \beta_k I_n)^{-1}(A - \alpha_k I_n), \quad \mathcal{B}_k := (B - \alpha_k I_m)^{-1}(B - \beta_k I_m), \\
\mathcal{F}_k &:= (A - \beta_k I_n)^{-1} F \in \mathbb{R}^{n \times r}, \quad \mathcal{G}_k := (B - \alpha_k I_m)^{-H} G \in \mathbb{C}^{m \times r}.
\end{aligned}$$

Now set $X_0 = 0$ and find factorization $X_k = Z_k D_k Y_k^H$

$$X_1 = \mathcal{A}_1 X_0 \mathcal{B}_1 + (\beta_1 - \alpha_1)\mathcal{F}_1 \mathcal{G}_1^H$$

,

LR-ADI Method

Sylvester ADI iteration [Wachspress 1988]

$$X_k = \mathcal{A}_k X_{k-1} \mathcal{B}_k + (\beta_k - \alpha_k)\mathcal{F}_k \mathcal{G}_k^H,$$
$$\mathcal{A}_k := (A - \beta_k I_n)^{-1}(A - \alpha_k I_n), \quad \mathcal{B}_k := (B - \alpha_k I_m)^{-1}(B - \beta_k I_m),$$
$$\mathcal{F}_k := (A - \beta_k I_n)^{-1} F \in \mathbb{R}^{n\times r}, \quad \mathcal{G}_k := (B - \alpha_k I_m)^{-H} G \in \mathbb{C}^{m\times r}.$$

Now set $X_0 = 0$ and find factorization $X_k = Z_k D_k Y_k^H$

$$X_1 = (\beta_1 - \alpha_1)(A - \beta_1 I_n)^{-1} F G^T (B - \alpha_1 I_m)^{-1}$$
$$\Rightarrow V_1 := Z_1 = (A - \beta_1 I_n)^{-1} F \in \mathbb{R}^{n\times r},$$
$$D_1 = (\beta_1 - \alpha_1) I_r \in \mathbb{R}^{r\times r},$$
$$W_1 := Y_1 = (B - \alpha_1 I_m)^{-H} G \in \mathbb{C}^{m\times r}.$$

LR-ADI Method

Sylvester ADI iteration [Wachspress 1988]

$$X_k = \mathcal{A}_k X_{k-1} \mathcal{B}_k + (\beta_k - \alpha_k)\mathcal{F}_k \mathcal{G}_k^H,$$
$$\mathcal{A}_k := (A - \beta_k I_n)^{-1}(A - \alpha_k I_n), \quad \mathcal{B}_k := (B - \alpha_k I_m)^{-1}(B - \beta_k I_m),$$
$$\mathcal{F}_k := (A - \beta_k I_n)^{-1} F \in \mathbb{R}^{n \times r}, \quad \mathcal{G}_k := (B - \alpha_k I_m)^{-H} G \in \mathbb{C}^{m \times r}.$$

Now set $X_0 = 0$ and find factorization $X_k = Z_k D_k Y_k^H$

$$\begin{aligned}
X_2 &= \mathcal{A}_2 X_1 \mathcal{B}_2 + (\beta_2 - \alpha_2)\mathcal{F}_2 \mathcal{G}_2^H \\
V_2 &= V_1 + (\beta_2 - \alpha_1)(A + \beta_2 I)^{-1} V_1 \in \mathbb{R}^{n \times r}, \\
W_2 &= W_1 + \overline{(\alpha_2 - \beta_1)}(B + \alpha_2 I)^{-H} W_1 \in \mathbb{R}^{m \times r}, \\
Z_2 &= [Z_1, \ V_2], \\
D_2 &= \operatorname{diag}(D_1, (\beta_2 - \alpha_2) I_r), \\
Y_2 &= [Y_1, \ W_2].
\end{aligned}$$

LR-ADI Algorithm [B. 2005, Li/Truhar 2008, B./Li/Truhar 2009]

Algorithm 1: Low-rank Sylvester ADI / factored ADI (fADI)

Input : Matrices defining $AX - XB = FG^T$ and shift parameters $\{\alpha_1, \ldots, \alpha_{k_{\max}}\}$, $\{\beta_1, \ldots, \beta_{k_{\max}}\}$.

Output: Z, D, Y such that $ZDY^H \approx X$.

$Z_1 = V_1 = (A - \beta_1 I_n)^{-1} F$,

$Y_1 = W_1 = (B - \alpha_1 I_m)^{-H} G$.

$D_1 = (\beta_1 - \alpha_1) I_r$

for $k = 2, \ldots, k_{\max}$ **do**

  $V_k = V_{k-1} + (\beta_k - \alpha_{k-1})(A - \beta_k I_n)^{-1} V_{k-1}$.

  $W_k = W_{k-1} + \overline{(\alpha_k - \beta_{k-1})}(B - \alpha_k I_n)^{-H} W_{k-1}$.

  Update solution factors

$$Z_k = [Z_{k-1}, V_k], \quad Y_k = [Y_{k-1}, W_k], \quad D_k = \operatorname{diag}\left(D_{k-1}, (\beta_k - \alpha_k) I_r\right).$$

Disadvantages of Low-Rank ADI as of 2012:

1. No efficient stopping criteria:
   - Difference in iterates $\rightsquigarrow$ norm of added columns/step: not reliable, stops often too late.
   - Residual is a full dense matrix, can not be calculated as such.
2. Requires complex arithmetic for real coefficients when complex shifts are used.
3. Expensive (only semi-automatic) set-up phase to precompute ADI shifts.

Disadvantages of Low-Rank ADI as of 2012:

1. No efficient stopping criteria:
   - Difference in iterates $\rightsquigarrow$ norm of added columns/step: not reliable, stops often too late.
   - Residual is a full dense matrix, can not be calculated as such.
2. Requires complex arithmetic for real coefficients when complex shifts are used.
3. Expensive (only semi-automatic) set-up phase to precompute ADI shifts.

None of these disadvantages exists as of today

LR-ADI Method

Disadvantages of Low-Rank ADI as of 2012:

1. No efficient stopping criteria:
   - Difference in iterates $\rightsquigarrow$ norm of added columns/step: not reliable, stops often too late.
   - Residual is a full dense matrix, can not be calculated as such.
2. Requires complex arithmetic for real coefficients when complex shifts are used.
3. Expensive (only semi-automatic) set-up phase to precompute ADI shifts.

None of these disadvantages exists as of today

Key observation: residual is low-rank matrix, with rank less than or equal to that of right-hand side!

$\Longrightarrow$ **speed-ups old vs. new LR-ADI can be up to 20!**

Other Low-rank Lyapunov Solvers...

...for Lyapunov equation $0 = AX + XA^T + BB^T$

Projection-based methods for Lyapunov equations with $A + A^T < 0$:

1. Compute orthonormal basis $\mathrm{range}(Z)$, $Z \in \mathbb{R}^{n\times r}$, for subspace $\mathcal{Z} \subset \mathbb{R}^n$, $\dim \mathcal{Z} = r$.
2. Set $\hat{A} := Z^T AZ$, $\hat{B} := Z^T B$.
3. Solve small-size Lyapunov equation $\hat{A}\hat{X} + \hat{X}\hat{A}^T + \hat{B}\hat{B}^T = 0$.
4. Use $X \approx Z\hat{X}Z^T$.

Other Low-rank Lyapunov Solvers...

... for Lyapunov equation $0 = AX + XA^T + BB^T$

Projection-based methods for Lyapunov equations with $A + A^T < 0$:

1. Compute orthonormal basis $\mathrm{range}(Z)$, $Z \in \mathbb{R}^{n \times r}$, for subspace $\mathcal{Z} \subset \mathbb{R}^n$, $\dim \mathcal{Z} = r$.
2. Set $\hat{A} := Z^T A Z$, $\hat{B} := Z^T B$.
3. Solve small-size Lyapunov equation $\hat{A}\hat{X} + \hat{X}\hat{A}^T + \hat{B}\hat{B}^T = 0$.
4. Use $X \approx Z\hat{X}Z^T$.

Examples:

- Krylov subspace methods, i.e., for $m = 1$:

$$\mathcal{Z} = \mathcal{K}(A, B, r) = \mathrm{span}\{B, AB, A^2B, \ldots, A^{r-1}B\}$$

  [Saad 1990, Jaimoukha/Kasenally 1994, Jbilou 2002–2008].

- Extended Krylov subspace method (EKSM) [Simoncini 2007],

$$\mathcal{Z} = \mathcal{K}(A, B, r) \cup \mathcal{K}(A^{-1}, B, r).$$

- Rational Krylov subspace methods (RKSM) [Druskin/Simoncini 2011].

The New LR-ADI Applied to Lyapunov Equations

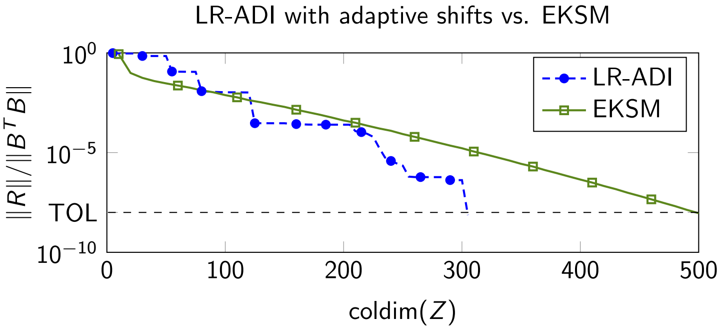
Example: an ocean circulation problem [Van Gijzen et al. 1998]

- FEM discretization of a simple 3D ocean circulation model (barotropic, constant depth) $\rightsquigarrow$ stiffness matrix $-A$ with $n = 42,249$, choose artificial constant term $B = \texttt{rand(n,5)}$.

The New LR-ADI Applied to Lyapunov Equations

Example: an ocean circulation problem [Van Gijzen et al. 1998]

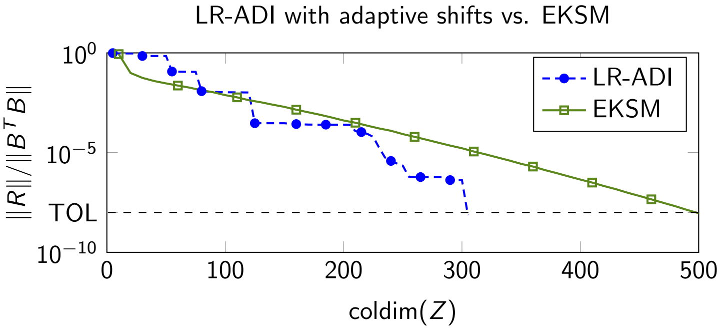
- FEM discretization of a simple 3D ocean circulation model (barotropic, constant depth) $\rightsquigarrow$ stiffness matrix $-A$ with $n = 42,249$, choose artificial constant term $B = \texttt{rand(n,5)}$.
- **Convergence history:**



The New LR-ADI Applied to Lyapunov Equations

Example: an ocean circulation problem [Van Gijzen et al. 1998]

- FEM discretization of a simple 3D ocean circulation model (barotropic, constant depth) $\rightsquigarrow$ stiffness matrix $-A$ with $n = 42,249$, choose artificial constant term $B =$ `rand(n,5)`.
- **Convergence history:**



- CPU times: LR-ADI $\approx$ 110 sec, EKSM $\approx$ 135 sec.

Summary & Outlook (Matrix Equations)

- Numerical enhancements of low-rank ADI for large Sylvester/Lyapunov equations:
  1. low-rank residuals, reformulated implementation,
  2. compute real low-rank factors in the presence of complex shifts,
  3. self-generating shift strategies (quantification in progress).

For diffusion-convection-reaction example:
332.02 sec. down to **17.24 sec.** $\rightsquigarrow$ acceleration by factor almost **20**.

- Generalized version enables derivation of low-rank solvers for various generalized Sylvester equations.
- Ongoing work:
  - Apply LR-ADI in Newton methods for algebraic Riccati equations

$$\mathcal{R}(X) = AX + XA^T + GG^T - XSS^TX = 0,$$

$$\mathcal{D}(X) = AXA^T - EXE^T + GG^T + A^TXF(I_r + F^TXF)^{-1}F^TXA = 0.$$

For nonlinear AREs see

P. Benner, P. Kürschner, J. Saak. *Low-rank Newton-ADI methods for large nonsymmetric algebraic Riccati equations.* J. Franklin Inst., 2015.

Outline

1. Motivation

2. Solving Large-Scale Sylvester and Lyapunov Equations

3. From Matrix Equations to PDEs in d Dimensions
   - The Curse of Dimensionality
   - Tensor Techniques
   - Numerical Examples

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.
$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\, x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \mathrm{vec}\,(X)$ and $b = \mathrm{vec}\,(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \to \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** $(d > 3)$.

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\, x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \operatorname{vec}(X)$ and $b = \operatorname{vec}(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

- Low-rankness of $\tilde{X} := VW^T \approx X$ follows from properties of A and B, e.g., [Penzl '00, Grasedyck '04].

From Matrix Equations to PDEs in d Dimensions

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\, x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \operatorname{vec}(X)$ and $b = \operatorname{vec}(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

- Low-rankness of $\tilde{X} := VW^T \approx X$ follows from properties of A and B, e.g.,
  [Penzl '00, Grasedyck '04].
- We solve this using low-rank Krylov subspace solvers. These essentially require matrix-vector multiplication and vector computations.

From Matrix Equations to PDEs in d Dimensions

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\, x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \operatorname{vec}(X)$ and $b = \operatorname{vec}(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

- Low-rankness of $\tilde{X} := VW^T \approx X$ follows from properties of A and B, e.g., [Penzl '00, Grasedyck '04].
- We solve this using low-rank Krylov subspace solvers. These essentially require matrix-vector multiplication and vector computations.
- Hence, $\mathcal{A}\operatorname{vec}(X_k) = \mathcal{A}\operatorname{vec}(V_k W_k^T) = \operatorname{vec}\left([\, AV_k,\ V_k\,]\,[\, W_k,\ AW_k\,]^T\right)$

From Matrix Equations to PDEs in d Dimensions

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\,x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \operatorname{vec}(X)$ and $b = \operatorname{vec}(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

- Low-rankness of $\tilde{X} := VW^T \approx X$ follows from properties of A and B, e.g., [Penzl '00, Grasedyck '04].
- We solve this using low-rank Krylov subspace solvers. These essentially require matrix-vector multiplication and vector computations.
- Hence, $\mathcal{A}\operatorname{vec}(X_k) = \mathcal{A}\operatorname{vec}(V_k W_k^T) = \operatorname{vec}\left([\,AV_k,\ V_k\,][\,W_k,\ AW_k\,]^T\right)$
- The rank of $[AV_k \quad V_k] \in \mathbb{R}^{n,2r}$, $[W_k \quad AW_k] \in \mathbb{R}^{n_t,2r}$ increases but can be controlled using truncation operator $\mathcal{T}_\varepsilon$ (like in sign function solver for Lyapunov equations). $\rightsquigarrow$ Low-rank Krylov subspace solvers. [Kressner/Tobler, B/Breiten, Savostyanov/Dolgov, ...].

From Matrix Equations to PDEs in d Dimensions

The Curse of Dimensionality [Bellman 1957]

Increase matrix size of discretized differential operator for $h \rightarrow \frac{h}{2}$ by factor 2^d.

$\rightsquigarrow$ **Rapid Increase of Dimensionality**, called **Curse of Dimensionality** ($d > 3$).

Consider $-\Delta u = f$ in $[0, 1] \times [0, 1] \subset \mathbb{R}^2$, uniformly discretized as

$$(I \otimes A + A \otimes I)\, x =: \mathcal{A}x = b \qquad \Longleftrightarrow \qquad AX + XA^T = B$$

with $x = \operatorname{vec}(X)$ and $b = \operatorname{vec}(B)$ with low-rank right hand side $B \approx b_1 b_2^T$.

- Low-rankness of $\tilde{X} := VW^T \approx X$ follows from properties of A and B, e.g., [Penzl '00, Grasedyck '04].
- We solve this using low-rank Krylov subspace solvers. These essentially require matrix-vector multiplication and vector computations.
- Hence, $\mathcal{A}\operatorname{vec}(X_k) = \mathcal{A}\operatorname{vec}(V_k W_k^T) = \operatorname{vec}\left([AV_k,\ V_k][W_k,\ AW_k]^T\right)$
- The rank of $[AV_k \quad V_k] \in \mathbb{R}^{n,2r}$, $[W_k \quad AW_k] \in \mathbb{R}^{n_t,2r}$ increases but can be controlled using truncation operator $\mathcal{T}_\varepsilon$ (like in sign function solver for Lyapunov equations). $\rightsquigarrow$ Low-rank Krylov subspace solvers. [Kressner/Tobler, B/Breiten, Savostyanov/Dolgov, ...].
- Ideas correspond to **separation of variables** in continuous-time: $b(x, y) \approx b_1(x) b_2(y)$.

From Matrix Equations to PDEs in d Dimensions

Ideas extend to $d \geq 3$: let $A_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, \ldots, d$, be 1d-discretization matrices corresponding to "grid points" $x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n_j)}$, then under certain assumptions, the Laplace operator (w/ homogeneous boundary conditions) can be discretized as

$$\mathcal{A} = A_1 \otimes I \otimes \cdots \otimes I + I \otimes A_2 \otimes I \otimes \cdots \otimes I + \ldots + I \otimes \cdots \otimes I \otimes A_d.$$

Note: if $n_j = n \, \forall \, j$, then $\mathcal{A} \in \mathbb{R}^{n^d \times n^d}$!

From Matrix Equations to PDEs in d Dimensions

Ideas extend to $d \geq 3$: let $A_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, \ldots, d$, be 1d-discretization matrices corresponding to "grid points" $x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n_j)}$, then under certain assumptions, the Laplace operator (w/ homogeneous boundary conditions) can be discretized as

$$\mathcal{A} = A_1 \otimes I \otimes \cdots \otimes I + I \otimes A_2 \otimes I \otimes \cdots \otimes I + \ldots + I \otimes \cdots \otimes I \otimes A_d.$$

Note: if $n_j = n \, \forall j$, then $\mathcal{A} \in \mathbb{R}^{n^d \times n^d}$!

If source term can be written as $f(x_1, x_2, \ldots, x_d) = f_1(x_1) f_2(x_2) \cdots f_d(x_d)$, discretized right-hand side becomes

$$b = b_1 \otimes b_2 \otimes \cdots \otimes b_d \in \mathbb{R}^{n^d} \quad \text{with} \quad b_j = [\, f_j(x_j^{(1)}), \ldots, f_j(x_j^{(n)}) \,].$$

From Matrix Equations to PDEs in d Dimensions

Ideas extend to $d \geq 3$: let $A_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, \ldots, d$, be 1d-discretization matrices corresponding to "grid points" $x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n_j)}$, then under certain assumptions, the Laplace operator (w/ homogeneous boundary conditions) can be discretized as

$$\mathcal{A} = A_1 \otimes I \otimes \cdots \otimes I + I \otimes A_2 \otimes I \otimes \cdots \otimes I + \ldots + I \otimes \cdots \otimes I \otimes A_d.$$

Note: if $n_j = n \; \forall \, j$, then $\mathcal{A} \in \mathbb{R}^{n^d \times n^d}$!

If source term can be written as $f(x_1, x_2, \ldots, x_d) = f_1(x_1) f_2(x_2) \cdots f_d(x_d)$, discretized right-hand side becomes

$$b = b_1 \otimes b_2 \otimes \cdots \otimes b_d \in \mathbb{R}^{n^d} \quad \text{with} \quad b_j = [\, f_j(x_j^{(1)}), \ldots, f_j(x_j^{(n)}) \,].$$

Now if the solution x of $\mathcal{A}x = b$

1. could be written as $x = x_1 \otimes x_2 \otimes \cdots \otimes x_d$ with $x_j \in \mathbb{R}^n$,
2. and could be computed without forming $\mathcal{A}, b$, but working only with A_j, b_j,

then we could hope to reduce the storage and computational complexity from n^d to dn!

From Matrix Equations to PDEs in d Dimensions

Ideas extend to $d \geq 3$: let $A_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, \ldots, d$, be 1d-discretization matrices corresponding to "grid points" $x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n_j)}$, then under certain assumptions, the Laplace operator (w/ homogeneous boundary conditions) can be discretized as

$$\mathcal{A} = A_1 \otimes I \otimes \cdots \otimes I + I \otimes A_2 \otimes I \otimes \cdots \otimes I + \ldots + I \otimes \cdots \otimes I \otimes A_d.$$

Note: if $n_j = n \, \forall \, j$, then $\mathcal{A} \in \mathbb{R}^{n^d \times n^d}$!

If source term can be written as $f(x_1, x_2, \ldots, x_d) = f_1(x_1) f_2(x_2) \cdots f_d(x_d)$, discretized right-hand side becomes

$$b = b_1 \otimes b_2 \otimes \cdots \otimes b_d \in \mathbb{R}^{n^d} \quad \text{with} \quad b_j = [\, f_j(x_j^{(1)}), \ldots, f_j(x_j^{(n)}) \,].$$

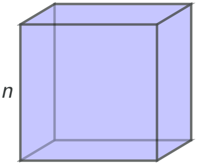
Now if the solution x of $\mathcal{A}x = b$

1. could be written as $x = x_1 \otimes x_2 \otimes \cdots \otimes x_d$ with $x_j \in \mathbb{R}^n$,
2. and could be computed without forming $\mathcal{A}, b$, but working only with A_j, b_j,

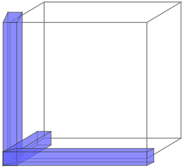
then we could hope to reduce the storage and computational complexity from n^d to dn! This requires **low-rank tensor techniques.**

(Note: for x we will need a more complex representation as 1. does not hold in general!)

Separation of Variables and Low-rank Approximation



$\rightsquigarrow$



- Approximate: $\underbrace{\mathbf{x}(i_1, \ldots, i_d)}_{\text{tensor}} \approx \underbrace{\sum_{\alpha} \mathbf{x}_{\alpha}^{(1)}(i_1)\mathbf{x}_{\alpha}^{(2)}(i_2) \cdots \mathbf{x}_{\alpha}^{(d)}(i_d)}_{\text{tensor product decomposition}}.$

Goals:

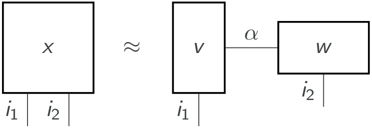
- Store and manipulate x — $\mathcal{O}(dn)$ cost *instead of* $\mathcal{O}(n^d)$.
- Solve equations $Ax = b$ — $\mathcal{O}(dn^2)$ cost *instead of* $\mathcal{O}(n^{2d})$.

Data Compression in 2D: Low-Rank Matrices

- Discrete *separation of variables*:

$$\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix} = \sum_{\alpha=1}^{r} \begin{bmatrix} v_{1,\alpha} \\ \vdots \\ v_{n,\alpha} \end{bmatrix} \begin{bmatrix} w_{\alpha,1} & \cdots & w_{\alpha,n} \end{bmatrix} + \mathcal{O}(\varepsilon).$$

- Diagrams:



- *Rank* $r \ll n$.
- $\mathtt{mem}(v) + \mathtt{mem}(w) = 2nr \ll n^2 = \mathtt{mem}(x)$.
- *Singular Value Decomposition (SVD)*
  $\Longrightarrow \quad \varepsilon(r)$ optimal w.r.t. spectral/Frobenius norm.

Tensor Trains

- Matrix Product States/Tensor Train (TT) format
  [Wilson '75, White '93, Verstraete '04, Oseledets '09/'11]:

  For indices

  $$\overline{i_p \ldots i_q} = (i_p - 1)n_{p+1} \cdots n_q + (i_{p+1} - 1)n_{p+2} \cdots n_q + \cdots + (i_{q-1} - 1)n_q + i_q,$$

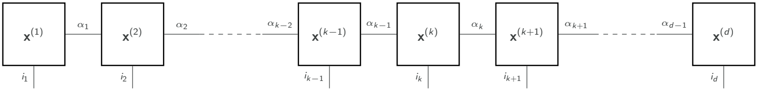
  the TT format can be expressed as

  $$x(\overline{i_1 \ldots i_d}) = \sum_{\boldsymbol{\alpha}=\mathbf{1}}^{\mathbf{r}} \mathbf{x}^{(1)}_{\alpha_1}(i_1) \cdot \mathbf{x}^{(2)}_{\alpha_1,\alpha_2}(i_2) \cdot \mathbf{x}^{(3)}_{\alpha_2,\alpha_3}(i_3) \cdots \mathbf{x}^{(d)}_{\alpha_{d-1},\alpha_d}(i_d)$$

  or

  $$x(\overline{i_1 \ldots i_d}) = \mathbf{x}^{(1)}(i_1) \cdots \mathbf{x}^{(d)}(i_d), \qquad \mathbf{x}^{(k)}(i_k) \in \mathbb{R}^{r_{k-1} \times r_k}.$$

  or



- Storage: $\mathcal{O}(dnr^2)$.

Overloading Tensor Operations

Always work with *factors* $\mathbf{x}^{(k)} \in \mathbb{R}^{r_{k-1} \times n_k \times r_k}$ instead of full tensors.

- Sum $z = x + y \quad \rightsquigarrow \quad$ increase of tensor rank $\quad r_z = r_x + r_y$.
- TT format for a high-dimensional operator

$$A(\overline{i_1 \ldots i_d}, \overline{j_1 \ldots j_d}) = \mathbf{A}^{(1)}(i_1, j_1) \cdots \mathbf{A}^{(d)}(i_d, j_d)$$

- *Matrix-vector* multiplication $y = Ax; \quad \rightsquigarrow$ tensor rank $r_y = r_A \cdot r_x$.
- Additions and multiplications *increase* TT ranks.
- *Decrease* ranks quasi-optimally via truncation operator $\mathcal{T}_\varepsilon$ using SVD (or QR).
- Iterative linear solvers (and preconditioners) can be implemented in TT format; we use, e.g., TT-GMRES.

Numerical Examples

Our work

Apply low-rank iterative solvers in TT format to discrete optimality systems resulting from

PDE-constrained optimization problems under uncertainty,

discretized by **stochastic Galerkin method**, i.e., applying Galerkin projection to weak formulation using

- implicit Euler (discontinuous Galerkin, dG(0)) in time,
- classical finite element methods (FEM) in physical space,
- generalized polynomial chaos in N-dimensional parameter space (resulting from parameterizing uncertain parameters).

This yields naturally a $(1 + 2(3) + N)$-way tensor representation of solution, which we approximate in low-rank TT format.

Numerical Examples

Our work

Apply low-rank iterative solvers in TT format to discrete optimality systems resulting from

PDE-constrained optimization problems under uncertainty,

discretized by **stochastic Galerkin method**.
This yields naturally a $(1 + 2(3) + N)$-way tensor representation of solution, which we approximate in low-rank TT format.

Take home message

Biggest problem solved so far has $n = 1.29 \cdot 10^{15}$ unknowns (optimality system for unsteady incompressible Navier-Stokes control problem with uncertain viscosity).

Numerical Examples

Our work

Apply low-rank iterative solvers in TT format to discrete optimality systems resulting from

PDE-constrained optimization problems under uncertainty,

discretized by **stochastic Galerkin method**

This yields naturally a $(1+2(3)+N)$-way tensor representation of solution, which we approximate in low-rank TT format.

Take home message

Biggest problem solved so far has $n = 1.29 \cdot 10^{15}$ unknowns (optimality system for unsteady incompressible Navier-Stokes control problem with uncertain viscosity).

Would require ≈ 10 **petabytes (PB)** $= 10,000$ **TB** to store the solution vector!

Numerical Examples

Our work

Apply low-rank iterative solvers in TT format to discrete optimality systems resulting from

PDE-constrained optimization problems under uncertainty,

discretized by **stochastic Galerkin method**

This yields naturally a $(1 + 2(3) + N)$-way tensor representation of solution, which we approximate in low-rank TT format.

Take home message

Biggest problem solved so far has $n = 1.29 \cdot 10^{15}$ unknowns (optimality system for unsteady incompressible Navier-Stokes control problem with uncertain viscosity).

Would require ≈ 10 **petabytes (PB)** $= 10,000$ **TB** to store the solution vector!

Using low-rank tensor techniques, we need $\approx 7 \cdot 10^7$ **bytes** $= 70$ **GB** to solve the optimality system in MATLAB in less than one hour!

Optimal Control of Unsteady Heat Equation

Consider the optimization problem

$$\mathcal{J}(t,y,u) = \frac{1}{2}||y-\bar{y}||^2_{L^2(0,T;\mathcal{D})\otimes L^2(\Omega)} + \frac{\alpha}{2}||\mathrm{std}(y)||^2_{L^2(0,T;\mathcal{D})} + \frac{\beta}{2}||u||^2_{L^2(0,T;\mathcal{D})\otimes L^2(\Omega)}$$

subject, $\mathbb{P}$-almost surely, to

$$\begin{cases} \dfrac{\partial y(t,\mathbf{x},\omega)}{\partial t} - \nabla\cdot(a(\mathbf{x},\omega)\nabla y(t,\mathbf{x},\omega)) = u(t,\mathbf{x},\omega), \text{ in } (0,T]\times\mathcal{D}\times\Omega, \\ \qquad y(t,\mathbf{x},\omega) = 0, \text{ on } (0,T]\times\partial\mathcal{D}\times\Omega, \\ \qquad y(0,\mathbf{x},\omega) = y_0, \text{ in } \mathcal{D}\times\Omega, \end{cases}$$

where

- for any $z:\mathcal{D}\times\Omega\to\mathbb{R}$, $z(\mathbf{x},\cdot)$ is a random variable defined on the complete probability space $(\Omega,\mathcal{F},\mathbb{P})$ for each $\mathbf{x}\in\mathcal{D}$,
- $\exists\ 0 < a_{\min} < a_{\max} < \infty$ s.t. $\mathbb{P}(\omega\in\Omega : a(x,\omega)\in[a_{\min},a_{\max}]\ \forall x\in\mathcal{D}) = 1$.

Weak formulation of the random PDE

Seek $y \in H^1\left(0, T; H_0^1(\mathcal{D}) \otimes L^2(\Omega)\right)$ such that, $\mathbb{P}$-almost surely,

$$\langle y_t, v \rangle + \mathcal{B}(y, v) = \ell(u, v) \quad \forall\, v \in H_0^1(\mathcal{D}) \otimes L^2(\Omega),$$

with the coercive[1] bilinear form

$$\mathcal{B}(y, v) := \int_\Omega \int_{\mathcal{D}} a(\mathbf{x}, \omega) \nabla y(\mathbf{x}, \omega) \cdot \nabla v(\mathbf{x}, \omega) d\mathbf{x}\, d\mathbb{P}(\omega), \quad v, y \in H_0^1(\mathcal{D}) \otimes L^2(\Omega),$$

and

$$\begin{aligned} \ell(u, v) &= \langle u(\mathbf{x}, \omega), v(\mathbf{x}, \omega) \rangle \\ &=: \int_\Omega \int_{\mathcal{D}} u(\mathbf{x}, \omega) v(\mathbf{x}, \omega) d\mathbf{x}\, d\mathbb{P}(\omega), \quad u, v \in H_0^1(\mathcal{D}) \otimes L^2(\Omega). \end{aligned}$$

Coercivity and boundedness of $\mathcal{B}$ + Lax-Milgram $\Longrightarrow$ unique solution exists.

[1] due to the positivity assumption on $a(\mathbf{x}, \omega)$

Weak formulation of the optimality system

Theorem [Chen/Quarteroni 2014, B./Onwunta/Stoll 2016]

Under appropriate regularity assumptions, there exists a unique adjoint state p and optimal solution (y, u, p) to the optimal control problem for the random unsteady heat equation, satisfying the stochastic optimality conditions (KKT system) for $t \in (0, T]$ $\mathbb{P}$-almost surely

$$
\begin{aligned}
\langle y_t, v \rangle + \mathcal{B}(y, v) &= \ell(u, v), && \forall\, v \in H_0^1(\mathcal{D}) \otimes L^2(\Omega),\\
\langle p_t, w \rangle - \mathbf{B}^*(p, w) &= \ell\left((y - \bar{y}) + \frac{\alpha}{2}\mathcal{S}(y), w\right), && \forall\, w \in H_0^1(\mathcal{D}) \otimes L^2(\Omega),\\
\ell(\beta u - p, \tilde{w}) &= 0, && \forall\, \tilde{w} \in L^2(\mathcal{D}) \otimes L^2(\Omega),
\end{aligned}
$$

where

- $\mathcal{S}(y)$ is the Fréchet derivative of $\|\mathrm{std}(y)\|^2_{L^2(0,T;\mathcal{D})}$;
- $\mathcal{B}^*$ is the adjoint operator of $\mathcal{B}$.

Discretization of the random PDE

- y, p, u are approximated using standard Galerkin ansatz, yielding approximations of the form

$$z(t, \mathbf{x}, \omega) = \sum_{k=0}^{P-1} \sum_{j=1}^{J} z_{jk}(t)\phi_j(\mathbf{x})\psi_k(\xi) = \sum_{k=0}^{P-1} z_k(t, \mathbf{x})\psi_k(\xi).$$

- Here,
  - $\{\phi_j\}_{j=1}^{J}$ are linear finite elements;
  - $\{\psi_k\}_{k=0}^{P-1}$ are the $P = \frac{(N+n)!}{N!n!}$ multivariate Legendre polynomials of degree $\leq n$.
  - Implicit Euler/dG(0) used for temporal discretization w/ constant time step τ.

The Fully Discretized Optimal Control Problem

Discrete first order optimality conditions/KKT system

$$\begin{bmatrix} \tau\mathcal{M}_1 & 0 & -\mathcal{K}_t^T \\ 0 & \beta\tau\mathcal{M}_2 & \tau\mathcal{N}^T \\ -\mathcal{K}_t & \tau\mathcal{N} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \tau\mathcal{M}_\alpha\bar{\mathbf{y}} \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix},$$

where

- $\mathcal{M}_1 = D \otimes G_\alpha \otimes M =: D \otimes \mathcal{M}_\alpha, \quad \mathcal{M}_2 = D \otimes G_0 \otimes M,$
- $\mathcal{K}_t = I_{n_t} \otimes \left[\sum_{i=0}^{N} G_i \otimes \widehat{K}_i\right] + (C \otimes G_0 \otimes M),$
- $\mathcal{N} = I_{n_t} \otimes G_0 \otimes M,$

and

- $G_0 = \operatorname{diag}\left(\langle\psi_0^2\rangle, \langle\psi_1^2\rangle, \ldots, \left\langle\psi_{P-1}^2\right\rangle\right), \quad G_i(j,k) = \langle\xi_i\psi_j\psi_k\rangle, \quad i = 1,\ldots,N,$
- $G_\alpha = G_0 + \alpha \operatorname{diag}\left(0, \langle\psi_1^2\rangle, \ldots, \left\langle\psi_{P-1}^2\right\rangle\right)$ (with first moments $\langle\,.\,\rangle$ w.r.t. $\mathbb{P}$),
- $\hat{K}_0 = M + \tau K_0, \quad \hat{K}_i = \tau K_i, \quad i = 1,\ldots,N,$
- $M, K_i \in \mathbb{R}^{J\times J}$ are the mass and stiffness matrices w.r.t. the spatial discretization, where K_i corresponds to the contributions of the ith KLE term to the stiffness,
- $C = -\operatorname{diag}(\text{ones}, -1), \quad D = \operatorname{diag}\left(\frac{1}{2}, 1, \ldots, 1, \frac{1}{2}\right) \in \mathbb{R}^{n_t\times n_t}.$

Discrete first order optimality conditions/KKT system

$$\begin{bmatrix} \tau\mathcal{M}_1 & 0 & -\mathcal{K}_t^T \\ 0 & \beta\tau\mathcal{M}_2 & \tau\mathcal{N}^T \\ -\mathcal{K}_t & \tau\mathcal{N} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \tau\mathcal{M}_\alpha \bar{\mathbf{y}} \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix},$$

Linear system with $3JPn_t$ unknowns!

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.
- Krylov subspace methods for indefinite symmetric systems: MINRES, ...

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.
- Krylov subspace methods for indefinite symmetric systems: MINRES, ...
- Requires good preconditioner.

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.
- Krylov subspace methods for indefinite symmetric systems: MINRES, ...
- Requires good preconditioner.
- Famous three-iterations-convergence result [Murphy/Golub/Wathen 2000]: using ideal preconditioner

  $$\mathcal{P} := \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix} \qquad \text{with the Schur complement} \quad S := -BA^{-1}B^T,$$

  MINRES finds the exact solution in at most three steps.

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.
- Krylov subspace methods for indefinite symmetric systems: MINRES, ...
- Requires good preconditioner.
- Famous three-iterations-convergence result [Murphy/Golub/Wathen 2000]: using ideal preconditioner

  $$\mathcal{P} := \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix} \qquad \text{with the Schur complement} \quad S := -BA^{-1}B^T,$$

  MINRES finds the exact solution in at most three steps.
- Motivates to use mean-based approximate Schur complement preconditioner

  $$\begin{bmatrix} \hat{A} & 0 \\ 0 & \hat{S} \end{bmatrix}.$$

Solving the Optimality System

Optimality system leads to saddle point problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Very large scale setting, (block-)structured sparsity $\rightsquigarrow$ iterative solution.
- Krylov subspace methods for indefinite symmetric systems: MINRES, ...
- Requires good preconditioner.
- Famous three-iterations-convergence result [Murphy/Golub/Wathen 2000]: using ideal preconditioner

  $$\mathcal{P} := \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix} \qquad \text{with the Schur complement} \quad S := -BA^{-1}B^T,$$

  MINRES finds the exact solution in at most three steps.
- Motivates to use mean-based approximate Schur complement preconditioner

  $$\begin{bmatrix} \hat{A} & 0 \\ 0 & \hat{S} \end{bmatrix}.$$
- Implemented in TT format $\rightsquigarrow$ TT-MINRES.

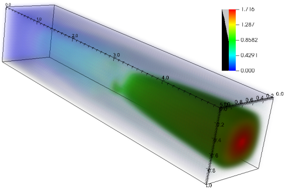
Mean-Based Preconditioned TT-MINRES

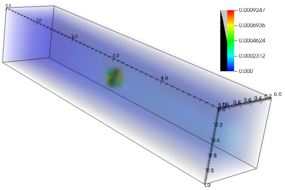
TT-MINRES	# iter (t)	# iter (t)	# iter (t)
n_t	2^5	2^6	2^8
$\dim(\mathcal{A}) = 3JPn_t$	$10,671,360$	$21,342,720$	$85,370,880$
$\alpha = 1,\ \mathrm{tol} = 10^{-3}$			
$\beta = 10^{-5}$	6 (285.5)	6 (300.0)	8 (372.2)
$\beta = 10^{-6}$	4 (77.6)	4 (130.9)	4 (126.7)
$\beta = 10^{-8}$	4 (56.7)	4 (59.4)	4 (64.9)
$\alpha = 0,\ \mathrm{tol} = 10^{-3}$			
$\beta = 10^{-5}$	4 (207.3)	6 (366.5)	6 (229.5)
$\beta = 10^{-6}$	4 (153.9)	4 (158.3)	4 (172.0)
$\beta = 10^{-8}$	2 (35.2)	2 (37.8)	2 (40.0)

Optimal Control 3D Stokes-Brinkman w/ Uncertain Viscosity

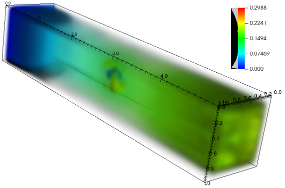
Mean | Standard deviation

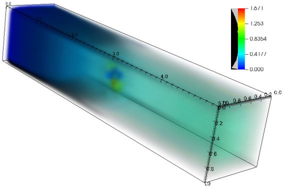
State





Control





- Full size: $n_{\mathbf{x}} n_{\xi} n_t \approx 3 \cdot 10^9$. Reduction: $\dfrac{\mathtt{mem}(TT)}{\mathtt{mem}(full)} = 0.002.$

Conclusions (High-dimensional PDEs)

- Low-rank tensor solver for unsteady heat **and Navier-Stokes equations** with uncertain viscosity.
- Similar techniques used for Stokes(-Brinkman) optimal control problems.
- Adapted AMEn (TT) solver to saddle point systems.
- To consider next:

Conclusions (High-dimensional PDEs)

- Low-rank tensor solver for unsteady heat **and Navier-Stokes equations** with uncertain viscosity.
- Similar techniques used for Stokes(-Brinkman) optimal control problems.
- Adapted AMEn (TT) solver to saddle point systems.
- To consider next:
  - Navier-Stokes: many parameters coming from uncertain geometry or Karhunen-Loève expansion of random fields;
    Initial results: the more parameters, the more significant is the complexity reduction w.r.t. memory — up to a factor of 10^9 for the control problem for a backward facing step.
  - exploit multicore technology (need efficient parallelization of AMEn).

References for Optimal Control under Uncertainty

P. Benner, S. Dolgov, A. Onwunta, and M. Stoll.
Solving optimal control problems governed by random Navier-Stokes equations using low-rank methods.
arXiv Preprint arXiv:1703.06097, March 2017.

P. Benner, S. Dolgov, A. Onwunta, and M. Stoll.
Low-rank solvers for unsteady Stokes-Brinkman optimal control problem with random data.
Computer Methods in Applied Mechanics and Engineering, 304:26–54, 2016.

P. Benner, A. Onwunta, and M. Stoll.
Low rank solution of unsteady diffusion equations with stochastic coefficients.
SIAM/ASA Journal on Uncertainty Quantification, 3(1):622–649, 2015.

P. Benner, A. Onwunta, and M. Stoll.
Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs.
SIAM Journal on Matrix Analysis and Applications, 37(2):491–518, 2016.

C.E. Powell and D.J. Silvester.
Preconditioning steady-state Navier–Stokes equations with random data.
SIAM Journal on Scientific Computing, 34(5):A2482–A2506, 2012.

E. Rosseel and G.N. Wells.
Optimal control with stochastic PDE constraints and uncertain controls.
Computer Methods in Applied Mechanics and Engineering, 213–216:152–167, 2012.

M. Stoll and T. Breiten.
A low-rank in time approach to PDE-constrained optimization.
SIAM Journal on Scientific Computing, 37(1):B1–B29, 2015.