(Randomized) Localized Model Order Reduction

Kathrin Smetana (University of Twente)

March 24, 2020

ICERM Workshop “Algorithms for Dimension and Complexity Reduction”
Motivation

- Model order reduction ...
  - ... allows to perform computations for many different configurations (parameters, geometry,...) very fast
  - ... without jeopardizing accuracy

- Topic of this talk: Localization and randomization facilitate (nearly) real-time simulations of large-scale problems
Outline

- Projection-based model order reduction in a nutshell
  - Randomized error estimation
- Localized Model Order Reduction
  - Constructing optimal local approximation spaces (in space)
  - Approximating optimal local approximation spaces via random sampling
  - Generating quasi-optimal local approximation spaces in time by random sampling
Parametrized Partial Differential Equation

- Parameter vector $\mu \in \mathcal{P}$; compact parameter set $\mathcal{P} \subset \mathbb{R}^P$
- **Parametrized PDE:** Given any $\mu \in \mathcal{P}$, find $u(\mu) \in X$, s.th.
  \[ A(\mu)u(\mu) = f(\mu) \quad \text{in } X'. \]

- $\Omega \subset \mathbb{R}^3$: bounded domain with Lipschitz boundary $\partial \Omega$
- $H^1_0(\Omega)^d \subset X \subset H^1(\Omega)^d \ (d = 1, 2, 3)$; $X'$: dual space
- $A(\mu) : X \to X'$: inf-sup stable, continuous linear differential operator
- $f(\mu) : X \to \mathbb{R}$: continuous linear form
Parametrized Partial Differential Equation

- Parameter vector $\mu \in \mathcal{P}$; compact parameter set $\mathcal{P} \subset \mathbb{R}^P$
- Parametrized PDE: Given any $\mu \in \mathcal{P}$, find $u(\mu) \in X$, s.th.
  \[
  A(\mu)u(\mu) = f(\mu) \quad \text{in } X'.
  \]

- High-dimensional discretization:
  - Introduce high-dimensional FE space $X^\mathcal{N} \subset X$ with $\dim(X^\mathcal{N}) = \mathcal{N}$ (assume small discretization error)
  - High-dimensional approximation: Given any $\mu \in \mathcal{P}$, find $u^\mathcal{N}(\mu) \in X^\mathcal{N}$, s.th.
    \[
    A(\mu)u^\mathcal{N}(\mu) = f(\mu) \quad \text{in } X^{\mathcal{N}'}.
    \]
- Issue: Require $u^\mathcal{N}(\mu)$ in real time and/or for many $\mu \in \mathcal{P}$. 

K Smetana (k.smetana@utwente.nl)
Parametrized Partial Differential Equation

- Parameter vector $\mu \in \mathcal{P}$; compact parameter set $\mathcal{P} \subset \mathbb{R}^P$
- Parametrized PDE: Given any $\mu \in \mathcal{P}$, find $u(\mu) \in X$, s.th.

$$A(\mu)u(\mu) = f(\mu) \quad \text{in } X'.$$

- High-dimensional discretization:
  - Introduce high-dimensional FE space $X^N \subset X$ with $\dim(X^N) = N$ (assume small discretization error)
  - High-dimensional approximation: Given any $\mu \in \mathcal{P}$, find $u^N(\mu) \in X^N$, s.th.

$$A(\mu)u^N(\mu) = f(\mu) \quad A(\mu) \in \mathbb{R}^{N \times N}, f(\mu) \in \mathbb{R}^N.$$

- Issue: Require $u^N(\mu)$ in real time and/or for many $\mu \in \mathcal{P}$. 
Projection-based model order reduction: key concept

- **Exploit**: $u^N(\mu)$ belongs to “solution manifold” $\mathcal{M}^N = \{u^N(\mu) \mid \mu \in \mathcal{P}\} \subset X^N$ of typically very low dimension

- **Offline**: Construct reduced space $X^N$ from solutions $u^N(\bar{\mu}_i)$, $i = 1, \ldots, N$ (e.g. by a Greedy algorithm, Proper Orthogonal Decomposition,...)

- **Online**: Galerkin projection on $X^N$: Given any $\mu^* \in \mathcal{P}$, find $u^N(\mu^*) \in X^N$, s.th.

\[
A(\mu^*) u^N(\mu^*) = f(\mu^*) \quad \text{in } (X^N)' \]

\[
\begin{bmatrix}
A(\mu)
\end{bmatrix}
\begin{bmatrix}
B
\end{bmatrix}
\begin{bmatrix}
u^N(\mu)
\end{bmatrix}
= 
\begin{bmatrix}
f(\mu)
\end{bmatrix}
\begin{bmatrix}
A^N(\mu)
\end{bmatrix}
\begin{bmatrix}
f^N(\mu)
\end{bmatrix}
\]
Construction of reduced basis $B$ via randomization

- **First Goal:** Given a matrix $S \in \mathbb{R}^{m \times n}$ and an integer $k$ find an orthonormal matrix $Q$ of rank $k$ such that $S \approx QQ^*S$.

- **Approach:**
  - Draw $k$ random vectors $r_j \in \mathbb{R}^n$ (say standard Gaussian)
  - Form sample vectors $y_j = Sr_j \in \mathbb{R}^m$ for $j = 1, \ldots, k$
  - Orthonormalize $y_j \rightarrow q_j$, $j = 1, \ldots, k$ and define $Q = [q_1, \ldots, q_k]$

- **Result:** If $S$ has exactly rank $k$ then $q_j$, $j = 1, \ldots, k$ span the range of $S$ at high probability. But also in the general case $q_j$, $j = 1, \ldots, k$ often perform nearly as good as the $k$ leading left singular vectors of $S$

- **Compute randomized SVD:**
  - Form $C = Q^*S$ which yields $S \approx QC$
  - Compute SVD of the small matrix $C = \tilde{U}\Sigma\tilde{V}^*$ and set $B = Q\tilde{U}$

For a review see for instance [Halko, Martinsson, Tropp 2011]
Construction of reduced basis \( B \) via randomization

- **First Goal:** Given a matrix \( S \in \mathbb{R}^{m \times n} \) and an integer \( k \) find an orthonormal matrix \( Q \) of rank \( k \) such that \( S \approx QQ^*S \).

- **Approach:**
  - Draw \( k \) random vectors \( r_j \in \mathbb{R}^n \) (say standard Gaussian)
  - Form sample vectors \( y_j = Sr_j \in \mathbb{R}^m \quad j = 1, \ldots, k \).
  - Orthonormalize \( y_j \mapsto q_j, \quad j = 1, \ldots, k \) and define \( Q = [q_1, \ldots, q_k] \)

- **Result:** If \( S \) has exactly rank \( k \) then \( q_j, \quad j = 1, \ldots, k \) span the range of \( S \) at high probability. But also in the general case \( q_j, \quad j = 1, \ldots, k \) often perform nearly as good as the \( k \) leading left singular vectors of \( S \)

- **Compute randomized SVD:**
  - Form \( C = Q^*S \) which yields \( S \approx QC \)
  - Compute SVD of the small matrix \( C = \tilde{U}\Sigma V^* \) and set \( B = Q\tilde{U} \)

Works also if \( S \) is not a data matrix but some linear map which is approximately low rank.
References for randomized construction of reduced models

- Hochman et al 2014
- Alla, Kutz 2015
- Zahm, Nouy 2016
- Balabanov, Nouy 2019, 2019
- Cohen, Dahmen, DeVore, Nichols 2020
- Saibaba 2020
A posteriori error estimation

- A posteriori error estimator is important both
  - to construct reduced order models via the greedy algorithm
  - to certify the approximation: how large is the error (in some QoI)?

**Proposition (A posteriori error bound)**

The error estimator $\tilde{\Delta}_N(\mu) = \beta_{LB}(\mu)^{-1}\|f(\mu) - A(\mu)u^N(\mu)\|_{X^N}$, with $\beta_{LB}(\mu) \leq \beta_N(\mu)$ satisfies

$$\|u^N(\mu) - u^N(\mu)\|_X \leq \tilde{\Delta}_N(\mu) \leq \frac{\gamma_N(\mu)}{\beta_{LB}(\mu)}\|u^N(\mu) - u^N(\mu)\|_X,$$

where $\beta_N(\mu) := \inf_{v \in X^N} \sup_{w \in X^N} \frac{\langle A(\mu)v, w \rangle}{\|v\|_X \|w\|_X}$ and $\gamma_N(\mu) = \sup_{v \in X^N} \sup_{w \in X^N} \frac{\langle A(\mu)v, w \rangle}{\|v\|_X \|w\|_X}$.

- Problem: Good estimate of stability constants often computationally infeasible; using simply the residual may perform very poorly, especially say for Helmholtz-type problems.
Outline

- Projection-based model order reduction in a nutshell
  - Randomized error estimation
- Localized Model Order Reduction
  - Constructing optimal local approximation spaces (in space)
  - Approximating optimal local approximation spaces via random sampling
  - Generating quasi-optimal local approximation spaces in time by random sampling

References:
References for randomization within error estimation

- Cao, Petzold 2004, Homescu, Petzold, Serban 2005
- Drohmann, Carlberg 2015, Trehan, Carlberg, and Durlofsky 2017
- Manzoni, Pagani, Lassila 2016
- Janon, Nodet, Prieur 2016
- Zahm, Nouy 2016
- Buhr, KS 2018
- Balabanov, Nouy 2019
- Eigel, Schneider, Trunschke, Wolf 2020
Randomized a posteriori error estimation

- **Goal:** Develop a posteriori error estimator for model order reduction that does not contain constants whose estimation is expensive (avoid estimating inf-sup constant and thus improve effectivity of estimator)

- **Setting:** We query a finite number of parameters for which we want to estimate the approximation error; allows computing statistics in UQ

- **Approach:** Exploit concentration inequalities:

**Proposition (Concentration inequality, Johnson-Lindenstrauss)**

Choose rows of matrix $\Phi \in \mathbb{R}^{K \times N}$ say as $K$ independent copies of standard Gaussian random vectors scaled by $1/\sqrt{K}$ and let $S \subset \mathbb{R}^N$ be a finite set. Moreover, assume $K \geq (C(z)/\varepsilon^2) \log(\#S/\delta)$. Then we have

$$\mathbb{P}\{(1 - \varepsilon)\|x - y\|_2^2 \leq \|\Phi x - \Phi y\|_2^2 \leq (1 + \varepsilon)\|x - y\|_2^2 \quad \forall x, y \in S\} \geq 1 - \delta.$$  

see for instance [Boucheron, Lugosi, Massart 2012], [Vershynin 2018]
Assumptions on random vector

- \( Z \in \mathbb{R}^N \): random vector such that

\[
\| v \|_\Sigma^2 = v^T \Sigma v = \mathbb{E}((Z^T v)^2) \quad \forall v \in \mathbb{R}^N,
\]

where \( \Sigma \) is matrix e.g. associated with \( H^1 \)- or \( L^2 \)-inner product or a quantity of interest

\( (Z^T v)^2 \) is an unbiased estimator of \( \| v \|_\Sigma^2 \)
Assumptions on random vector

- \( Z \in \mathbb{R}^N \): random vector such that

\[
\| v \|_\Sigma^2 = v^T \Sigma v = \mathbb{E}((Z^T v)^2) \quad \forall v \in \mathbb{R}^N,
\]

where \( \Sigma \) is matrix e.g. associated with \( H^1 \)- or \( L^2 \)-inner product or a quantity of interest

\( (Z^T v)^2 \) is an unbiased estimator of \( \| v \|_\Sigma^2 \)

- For simplicity: Assume \( Z \sim \mathcal{N}(0, \Sigma) \) is a Gaussian vector with zero mean and covariance matrix \( \Sigma \)
Assumptions on random vector

- $Z \in \mathbb{R}^N$: random vector such that

$$
\|v\|_\Sigma^2 = v^T \Sigma v = \mathbb{E}((Z^T v)^2) \quad \forall v \in \mathbb{R}^N,
$$

where $\Sigma$ is matrix e.g. associated with $H^1$- or $L^2$-inner product or a quantity of interest

$(Z^T v)^2$ is an unbiased estimator of $\|v\|_\Sigma^2$

- For simplicity: Assume $Z \sim \mathcal{N}(0, \Sigma)$ is a Gaussian vector with zero mean and covariance matrix $\Sigma$

- $Z_1, \ldots, Z_K$: $K$ independent copies of $Z$

- Consider the following (unbiased) Monte-Carlo estimator of $\|v\|_\Sigma^2$

$$
\frac{1}{K} \sum_{i=1}^{K} (Z_i^T v)^2.
$$
Proposition (Concentration inequality (KS, Zahm, Patera 2019))

Given a finite set of parameters $S = \{\mu_1, \ldots, \mu_S\} \subset \mathcal{P}$, a failure probability $0 < \delta < 1$, $w \in \mathbb{R}$, $w > \sqrt{e}$, we have for

$$K \geq \frac{\log(\#S) + \log(\delta^{-1})}{\log(w/\sqrt{e})}$$

that

$$\mathbb{P}\left\{ \frac{\| e(\mu_j) \|_\Sigma^2}{w^2} \leq \frac{1}{K} \sum_{i=1}^{K} (Z_i^T e(\mu_j))^2 \leq w^2 \| e(\mu_j) \|_\Sigma^2, \quad \forall \mu_j \in S \right\} \geq 1 - \delta.$$

- chi-squared distribution
- concentration around 1 (that means error estimator has close to perfect effectivity 1)
Proposition (Concentration inequality (KS, Zahm, Patera 2019))

Given a finite set of parameters $S = \{\mu_1, \ldots, \mu_S\} \subset \mathcal{P}$, a failure probability $0 < \delta < 1$, $w \in \mathbb{R}$, $w > \sqrt{e}$, we have for

$$K \geq \frac{\log(\#S) + \log(\delta^{-1})}{\log(w/\sqrt{e})}$$

that

$$\mathbb{P} \left\{ \frac{\|e(\mu_j)\|^2_\Sigma}{w^2} \leq \frac{1}{K} \sum_{i=1}^{K} (Z_i^T e(\mu_j))^2 \leq w^2\|e(\mu_j)\|^2_\Sigma, \quad \forall \mu_j \in S \right\} \geq 1 - \delta.$$

<table>
<thead>
<tr>
<th>$#S$</th>
<th>$w = 2$</th>
<th>$w = 3$</th>
<th>$w = 4$</th>
<th>$w = 5$</th>
<th>$w = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>100</td>
<td>48</td>
<td>16</td>
<td>11</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>1000</td>
<td>60</td>
<td>20</td>
<td>13</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>$10^6$</td>
<td>96</td>
<td>31</td>
<td>21</td>
<td>17</td>
<td>11</td>
</tr>
</tbody>
</table>

Table: Values for $K$ that guarantee (1) for all $\mu_j \in S$ with $\delta = 10^{-2}$. 
Proposition (Concentration inequality (KS, Zahm, Patera 2019))

Given a finite set of parameters $S = \{\mu_1, \ldots, \mu_S\} \subset \mathcal{P}$, a failure probability $0 < \delta < 1$, $w \in \mathbb{R}$, $w > \sqrt{e}$, we have for

$$K \geq \frac{\log(\#S) + \log(\delta^{-1})}{\log(w/\sqrt{e})}$$

that

$$\mathbb{P} \left\{ \frac{\|e(\mu_j)\|_\Sigma^2}{w^2} \leq \frac{1}{K} \sum_{i=1}^{K} (Z_i^T e(\mu_j))^2 \leq w^2 \|e(\mu_j)\|_\Sigma^2, \quad \forall \mu_j \in S \right\} \geq 1 - \delta.$$

Define $\Delta(\mu) := \left( \frac{1}{K} \sum_{i=1}^{K} (Z_i^T e(\mu))^2 \right)^{1/2}$

Problem: estimator $\Delta(\mu) = \left( \frac{1}{K} \sum_{i=1}^{K} (Z_i^T (u^N(\mu_j) - u^N(\mu_j)))^2 \right)^{1/2}$ involves high-dimensional finite element solution

$\implies$ Computationally infeasible in the online stage
A fast-to-evaluate randomized error estimator

- Exploit error residual relationship

\[ Z_i^T e(\mu) = Z_i^T A(\mu)^{-1} (f(\mu) - A(\mu) u^N(\mu)) = (A(\mu)^{-T} Z_i)^T r(\mu) \]

- Define solutions of dual problems with random right-hand sides \( Z_i \):

\[ Y^N_i(\mu) := A(\mu)^{-T} Z_i \]

- Approximation of the dual solutions via model order reduction:

\[ y^N_i(\mu) \approx y^{N_{du}}_i(\mu) \in \tilde{Y} \subset X^N; \quad \tilde{Y} : \text{dual reduced space.} \]

- Define fast-to-evaluate randomized error estimator

\[ \Delta^{N_{du}}(\mu) := \left( \frac{1}{K} \sum_{i=1}^{K} \left( y^{N_{du}}_i(\mu) T r(\mu) \right)^2 \right)^{1/2} \]
**Proposition**

*Choose $S \in \mathbb{N}$ in the offline stage. Then, in the online stage for any given $w > \sqrt{e}$ and $\delta > 0$ we have for $S$ different parameters values $\mu_j, j = 1, \ldots, S$ in a finite parameter set $S = \{\mu_1, \ldots, \mu_S\}$ and *

$$K \geq \frac{\log(S) + \log(\delta^{-1})}{\log(w/\sqrt{e})} \quad \text{that} \quad \Delta_{Ndu}^{N_{du}}(\mu_j) := \left(\frac{1}{K} \sum_{i=1}^{K} (y_i^{N_{du}}(\mu_j)^T r(\mu_j))^2\right)^{1/2}$$

satisfies

$$\mathbb{P}\left\{ (\alpha w)^{-1} \Delta_{Ndu}^{N_{du}}(\mu_j) \leq \| e(\mu_j) \|_\Sigma \leq (\alpha w) \Delta_{Ndu}^{N_{du}}(\mu_j), \quad \mu_j \in S, \right\} \geq 1 - \delta,$$

where

$$\alpha = \max_{\mu \in \mathcal{P}} \left( \max \left\{ \frac{\Delta(\mu)}{\Delta_{Ndu}(\mu)}, \frac{\Delta_{Ndu}(\mu)}{\Delta(\mu)} \right\} \right) \geq 1.$$
Numerical experiments: acoustics in 2D

- Consider on $\Omega = (0, 1) \times (0, 1)$

\[
-\partial_{x_1 x_1} u(x; \mu) - \mu_1 \partial_{x_2 x_2} u(x; \mu) - \mu_2 u(x; \mu) = f(x) \quad \text{in } \Omega,
\]

\[
u(x; \mu) = 0 \quad \text{on the bottom},
\]

\[\nabla u(x; \mu) \cdot n = 0 \quad \text{on the sides},
\]

\[
\kappa(\mu_1) \nabla u(x; \mu) \cdot n = \cos(\pi x) \quad \text{on the top}.
\]

- $m \in \mathcal{P} = [0.2, 1.2] \times [10, 50]$

- high dimensional discretization: linear FE, $h = 0.01$ in each direction
Randomized a posteriori error estimation

Numerical Experiments

Histograms of effectivity $\Delta_{N_{du}}^N u_N(\mu) - u^N(\mu) \| H^1(\Omega)$

Figure: $\#S = 10^4$, $N_{primal} = 20$, $q = 0.99$, 100 realizations, vertical dashed lines: $1/w$ and $w$, grey area: $1/(tol w)$ and $tol w$, where $\alpha \approx tol$, solid lines: chi-squared distribution
Outline

- Projection-based model order reduction in a nutshell
  - Randomized error estimation
- Localized Model Order Reduction
  - Constructing optimal local approximation spaces (in space)
  - Approximating optimal local approximation spaces via random sampling
  - Generating quasi-optimal local approximation spaces in time by random sampling

References:

Localized model order reduction

Limitations of standard model order reduction approach:

- **Curse of parameter dimensionality**: many parameters require prohibitively large reduced spaces
- **No topological flexibility** (although geometric variation is possible)
- Possibly **high computational costs in the offline stage**
Localized model order reduction

Limitations of standard model order reduction approach:

- **Curse of parameter dimensionality**: many parameters require prohibitively large reduced spaces
- **No topological flexibility** (although geometric variation is possible)
- Possibly high computational costs in the offline stage

→ **Localized model order reduction**

Further advantages:

- Allows to use different (sizes of) reduced spaces in different parts of the domain (similar to hp-methods)
- (Local) changes of the PDE, the geometry in the online stage are possible
Construction of local reduced spaces, some references

- Existing approaches ...
  - ... either provided a fast convergence but error analysis seems challenging: [Eftang, Patera 13], [Martini, Rozza, Haasdonk 15], ...
  - ... or came with a rigorous error analysis but slow convergence: [Hetmaniuk, Lehoucq 10], [Jakobsson, Bengzon, Larson 11], [Hetmaniuk, Klawonn 14], ...
Construction of local reduced spaces, some references

- **Existing approaches ...**
  - ... either provided a fast convergence but error analysis seems challenging: [Eftang, Patera 13], [Martini, Rozza, Haasdonk 15], ...
  - ... or came with a rigorous error analysis but slow convergence: [Hetmaniuk, Lehoucq 10], [Jakobsson, Bengzon, Larson 11], [Hetmaniuk, Klawonn 14], ...

- **Idea:** Use concepts from multiscale methods introduced in [Babuška, Lipton 11], [Malqvist, Peterseim 14] that ...
  - ... rely on the decay behavior of the solution of certain PDEs even for rough coefficients
  - ... and the compactness of certain operators thanks to the Caccioppoli inequality (bounds energy norm of solutions of the PDE by $L^2$-norm on a larger domain)

$\implies$ Yields **superalgebraic convergence and rigorous error analysis**
Localized model order reduction

Challenges:

- We can only exploit that the **global solution** solves PDE locally
- But: **No knowledge** of the trace of the global solution on $\Gamma_{\text{out}}$

$\implies$ Infinite dimensional parameter space
Localized model order reduction

Challenges:

- We can only exploit that the global solution solves PDE locally
- But: No knowledge of the trace of the global solution on $\Gamma_{out}$

$\Rightarrow$ Infinite dimensional parameter space

Idea:

- Restrict to space of functions that solve the PDE locally on $\Omega$ for arbitrary boundary conditions on $\Gamma_{out}$
- Exploit that for those local solutions we have a very fast decay of higher frequencies from $\Gamma_{out}$ to $\Omega_{in}, \Gamma_{in}$ ($\rightarrow$ Caccioppoli inequality)
- yields optimal local approximation spaces in the sense of Kolmogorov
Optimal local approximation spaces

Definition (Kolmogorov n-width, optimal subspaces (Kolmogoroff 1936))

$S$, $R$ Hilbert spaces, $R^n$: subspace of $R$, $\dim R^n = n$, $T : S \to R$ linear, continuous operator. The Kolmogorov $n$-width is defined as

$$d_n(T(S); R) := \inf_{\dim R^n = n} \sup_{\eta \in S} \inf_{\zeta \in R^n} \frac{\|T(\eta) - \zeta\|_R}{\|\eta\|_S}$$

A subspace $R^n$ with $\dim R^n \leq n$, that satisfies

$$d_n(T(S); R) = \sup_{\eta \in S} \inf_{\zeta \in R^n} \frac{\|T(\eta) - \zeta\|_R}{\|\eta\|_S}$$

is called an optimal subspace.
Motivation: separation of variables

- Consider $\Omega = (-5, 5) \times (0, 1)$
  
  $$-\Delta u = 0, \quad \text{in } \Omega, \quad \frac{du}{dy}(x, 1) = \frac{du}{dy}(x, 0) = 0.$$ 

- plus: arbitrary Dirichlet boundary conditions on $\Gamma_{out}$. 

\[ \begin{array}{c|c|c} 
\Gamma_{out} & \Gamma_{in} & \Gamma_{out} \\
\end{array} \]
Motivation: separation of variables

- Consider $\Omega = (-5, 5) \times (0, 1)$

\[-\Delta u = 0, \quad \text{in } \Omega, \quad \frac{du}{dy}(x, 1) = \frac{du}{dy}(x, 0) = 0.\]

- plus: arbitrary Dirichlet boundary conditions on $\Gamma_{out}$.
- separation of variables: all harmonic functions on $\Omega$ have the form

\[u(x, y) = a_0 + b_0 x + \sum_{n=1}^{\infty} \cos(n\pi y)[a_n \cosh(n\pi x) + b_n \sinh(n\pi x)].\]

- Example: Prescribe $\cos(3\pi y)$ on $\Gamma_{out}$ and thus $n = 3$:
Motivation: separation of variables

- Consider $\Omega = (-5, 5) \times (0, 1)$

$$-\Delta u = 0, \quad \text{in} \quad \Omega, \quad \frac{du}{dy}(x, 1) = \frac{du}{dy}(x, 0) = 0.$$ 

- plus: arbitrary Dirichlet boundary conditions on $\Gamma_{out}$.

- separation of variables: all harmonic functions on $\Omega$ have the form

$$u(x, y) = a_0 + b_0 x + \sum_{n=1}^{\infty} \cos(n\pi y) [a_n \cosh(n\pi x) + b_n \sinh(n\pi x)]$$

$\implies$ Extremely rapid and exponential decay of the cos-functions in the interior of $\Omega$ for higher $n$.

$\implies$ Most harmonic extensions of the basis functions $\cos(n\pi y)$, $n = 0, \ldots, \infty$ are practically zero on $\Gamma_{in}$.

$\implies$ A reduced space of very low dimension on $\Gamma_{in}$ will already yield a very good approximation!
Motivation: separation of variables

- Consider $\Omega = (-5, 5) \times (0, 1)$

  $$-\Delta u = 0, \quad \text{in } \Omega, \quad \frac{du}{dy}(x, 1) = \frac{du}{dy}(x, 0) = 0.$$ 

- plus: arbitrary Dirichlet boundary conditions on $\Gamma_{out}$.
- separation of variables: all harmonic functions on $\Omega$ have the form

  $$u(x, y) = a_0 + b_0x + \sum_{n=1}^{\infty} \cos(n\pi y)[a_n \cosh(n\pi x) + b_n \sinh(n\pi x)]$$

  $\implies$ Extremely rapid and exponential decay of the cos-functions in the interior of $\Omega$ for higher $n$.

Question: How can we generalize this idea?
The space of all local solutions of the PDE on $\Omega$

- Consider the space of all local solutions of the PDE\(^1\) on $\Omega$
  \[\mathcal{H} := \{ w \in H^1(\Omega) : \text{with } Aw = 0 \in X' \} .\]

- Global solution of the PDE restricted to $\Omega$ lies in $\mathcal{H}$!

- We are interested in $u|_{\Gamma_{in}}$ or $u|_{\Omega_{in}}$ and thus introduce
  \[ R := \{ w|_{\Gamma_{in}} , \ w \in \mathcal{H} \} \quad \text{or} \quad R := \{ w|_{\Omega_{in}} , \ w \in \mathcal{H} \}, \]
  and $S := \{ w|_{\Gamma_{out}} , \ w \in \mathcal{H} \}$.

\(^1\)For theoretical purposes one needs to consider the quotient space $\tilde{\mathcal{H}} := \mathcal{H}/\ker(A)$ at certain instances.
Transfer operator

- We introduce a transfer operator
  \[ T : S \rightarrow R \]

- For \( w \in \mathcal{H} \) and thus \( w|_{\Gamma_{out}} \in S \) we define
  \[ T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}. \]
Transfer operator

- We introduce a transfer operator

\[ T : S \rightarrow R \]

- For \( w \in \mathcal{H} \) and thus \( w|_{\Gamma_{out}} \in S \) we define

\[ T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}. \]
Transfer operator

- We introduce a transfer operator

\[ T : S \rightarrow R \]

- For \( w \in \mathcal{H} \) and thus \( w|_{\Gamma_{out}} \in S \) we define

\[ T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}. \]
Transfer operator

- We introduce a transfer operator $T : S \rightarrow R$
- For $w|_{\Gamma_{\text{out}}} \in S$ we define $T(w|_{\Gamma_{\text{out}}}) := w|_{\Gamma_{\text{in}}}$ or $T(w|_{\Gamma_{\text{out}}}) := w|_{\Omega_{\text{in}}}$.
- $T$ is compact thanks to the Caccioppoli inequality:

**Lemma (Caccioppoli inequality for heat conduction)**

Let $\kappa \in L^\infty(\Omega)$ fulfill $0 < \kappa_0 \leq \kappa \leq \kappa_1$ with constants $\kappa_0, \kappa_1$, define $X^0 = \{v \in H^1(\Omega), v|_{\Gamma_{\text{out}}} = 0\}$, let $u \in X := \{v \in H^1(\Omega), v|_{\Gamma_{\text{out}}} = g\}$ satisfy

$$\int_{\Omega} \kappa \nabla u \cdot \nabla v = 0 \quad \forall v \in X^0.$$

Then on $\Omega^* \subsetneq \Omega^{**} \subset \Omega$ with $\text{dist}(\partial \Omega^* \backslash \partial \Omega, \partial \Omega^{**} \backslash \partial \Omega) > \varrho > 0$ there holds

$$\int_{\Omega^*} \kappa |\nabla u|^2 \, dx \leq \frac{c}{\varrho^2} \|u\|_{L^2(\Omega^{**} \backslash \Omega^*)}^2.$$
Transfer operator

- We introduce a transfer operator $T : S \rightarrow R$
- For $w|_{\Gamma_{out}} \in S$ we define $T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}}$ or $T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}$.
- $T$ is compact thanks to the Caccioppoli inequality.
- Introduce adjoint operator $T^*$ and consider the eigenvalue problem

  $$T^* T w|_{out} = \lambda w|_{out} \quad \text{for } w \in \mathcal{H}.$$ 

- Equivalent formulation: Find $(\varphi_j, \lambda_j) \in (\mathcal{H}, \mathbb{R}^+) \text{ such that}$

  $$(\varphi_j|_{D_{in}}, w|_{D_{in}})_R = \lambda_j (\varphi_j|_{\Gamma_{out}}, w|_{\Gamma_{out}})_S \quad \forall w \in \mathcal{H}, D_{in} = \Gamma_{in}, \Omega_{in}$$
Transfer eigenvalue problem

Proposition (Transfer eigenvalue problem)

- \( \varphi_j \) and \( \lambda_j \): eigenfunctions and eigenvalues of the transfer eigenvalue problem: Find \( (\varphi_j, \lambda_j) \in (\mathcal{H}, \mathbb{R}^+) \) such that

\[
(\varphi_j|_{D_{\text{in}}}, w|_{D_{\text{in}}})_R = \lambda_j (\varphi_j|_{\Gamma_{\text{out}}}, w|_{\Gamma_{\text{out}}})_S \quad \forall w \in \mathcal{H}, D_{\text{in}} = \Gamma_{\text{in}}, \Omega_{\text{in}}
\]

- List \( \lambda_j \) such that \( \lambda_1 \geq \lambda_2 \geq \ldots \), and \( \lambda_j \to 0 \) as \( j \to \infty \).

- The optimal space on \( \Gamma_{\text{in}} \) or \( \Omega_{\text{in}} \) is given by

\[
R^n := \text{span}\{\phi_{1,sp}, \ldots, \phi_{n,sp}\}, \quad \phi_{j,sp} = T\varphi_j|_{\Gamma_{\text{out}}}, \quad j = 1, \ldots, n.
\]

- 

\[
d_n(T(S); R) = \sup_{\xi \in S} \inf_{\zeta \in R^n} \frac{\|T\xi - \zeta\|_R}{\|\xi\|_S} = \sqrt{\lambda_{n+1}}
\]
A priori error bound

Proposition (A priori error bound (KS, Patera 2016))

\[ |u - u^n| \leq C_1(\Omega) \sqrt{\lambda_{n+1}}, \]

where \( C_1(\Omega) \) does neither depend on \( u \) nor on \( u^n \).
Numerical experiments for isotropic linear elasticity

cracked I-Beam, uniform Young’s modulus $E_i = 1$ in both components

Figure: eigenvalues $\lambda_n$

Figure: component mesh
Numerical experiments for isotropic linear elasticity

Stiffened plate — simplified model for ship stiffener

- $E_i = 1$ in grey areas, $i = 1, 2$
- $E_i = E_i' \in [1, 20]$ varies in red areas

Figure: eigenvalues $\lambda_j$

Figure: mesh in $\Omega_i$
Numerical experiments for isotropic linear elasticity

Stiffened plate — simplified model for ship stiffener

Figure: eigenvalues $\lambda_j$

Figure: stiffened plate under bending
Comparison with other reduced interface spaces

Solid beam, \( E_i = E^r_i = 1^2, g|_{\Gamma_1} = (0, 0, 0)^T, g|_{\Gamma_2} = (1, 1, 1)^T \)

- Legendre polynomials: components of the displacement are solutions of scalar singular Sturm-Liouville problems
- Empirical port modes constructed by a pairwise training algorithm [Eftang, Patera 2013]
- Spectral modes constructed by the spectral greedy

\[ \mathcal{P}_i = [1, 10] \times [1, 1] \text{ for } \mu_i = (E_i, E^r_i) \]
Numerical experiments: shiploader

\[\text{Field: Von Mises (MPa)}\]

\[\begin{array}{c}
200 \\
171.4 \\
142.9 \\
114.3 \\
85.71 \\
57.14 \\
28.57 \\
0 
\end{array}\]

3 Results by company Akselos S.A.; KS has no financial interest in Akselos S.A.
Numerical experiments: shiploader

- Discretization with FEM: >20 millions of DOFs
- Size of Schur complement system: \( \approx 349 \, 000 \)

- Size of reduced Schur complement system: \( \approx 12 \, 000 \)
- Simulation time with reduced port spaces: \( \approx 2 \, \text{sec} \)

\(^3\text{Results by company Akselos S.A.; KS has no financial interest in Akselos S.A.}\)
Computing an approximation of the transfer eigenvalue problem

Transfer eigenvalue problem: Find \((\varphi_j, \lambda_j) \in (\mathcal{H}, \mathbb{R}^+)\) such that

\[
( T^h(\varphi_j|_{\Gamma_{out}}), T^h(w|_{\Gamma_{out}}) )_R = \lambda_j ( \varphi_j|_{\Gamma_{out}}, w|_{\Gamma_{out}} )_S \quad \forall w \in \mathcal{H}
\]

\[\mathcal{H} = \{ \text{set of all local solutions of the PDE with arbitrary Dirichlet b. c.} \}\]

1. Introduce a FE discretization with \(N_{out}\) degrees of freedom (DOFs) on \(\Gamma_{out}\) and \(N_{in}\) DOFs on \(\Gamma_{in}\) or \(\Omega_{in}\)
2. Solve for each basis function on \(\Gamma_{out}\) the PDE locally
   \(\implies\) number of required local solutions of the PDE scales with the number of DOFs on \(\Gamma_{out}\) and thus depends on the discretization
3. Assemble and solve generalized eigenvalue problem
Computing an approximation of the transfer eigenvalue problem

Transfer eigenvalue problem: Find \( (\varphi_j, \lambda_j) \in (\mathcal{H}, \mathbb{R}^+) \) such that

\[
( T^h(\varphi_j|_{\Gamma_{\text{out}}}), T^h(w|_{\Gamma_{\text{out}}}) )_R = \lambda_j ( \varphi_j|_{\Gamma_{\text{out}}}, w|_{\Gamma_{\text{out}}})_S \quad \forall w \in \mathcal{H}
\]

\( \mathcal{H} = \{ \text{set of all local solutions of the PDE with arbitrary Dirichlet b. c.} \} \)

1. Introduce a FE discretization with \( N_{\text{out}} \) degrees of freedom (DOFs) on \( \Gamma_{\text{out}} \) and \( N_{\text{in}} \) DOFs on \( \Gamma_{\text{in}} \) or \( \Omega_{\text{in}} \)
2. Solve for each basis function on \( \Gamma_{\text{out}} \) the PDE locally
   \( \quad \longrightarrow \) number of required local solutions of the PDE scales with the number of DOFs on \( \Gamma_{\text{out}} \) and thus depends on the discretization
3. Assemble and solve generalized eigenvalue problem

Problem: For large number of DOFs on \( \Gamma_{\text{out}} \) the approximation of the transfer eigenvalue problem can be very/prohibitively expensive especially in 3D
Outline

- Projection-based model order reduction in a nutshell
  - Randomized error estimation

- Localized Model Order Reduction
  - Constructing optimal local approximation spaces (in space)
  - Approximating optimal local approximation spaces via random sampling
  - Generating quasi-optimal local approximation spaces in time by random sampling

References on randomization in multiscale, domain decomposition methods

- Wang, Vouvakis 2015
- Calo, Efendiev, Galvis, Li 2016
- Owhadi 2015, 2017
Approximating optimal local spaces with Randomized Linear Algebra

- Prescribe random boundary conditions; in detail choose every coefficient of a FEM basis function on $\Gamma_{out}$ as a (mutually independent) Gaussian random variable with zero mean and variance one.
- Solve PDE for random boundary conditions numerically and store evaluation of local solution of PDE $u^h|_{\Gamma_{in}}$ or $u^h|_{\Omega_{in}}$.
- Define reduced space $R^n_{rand}$ as the span of $n$ such evaluations $u^h|_{\Gamma_{in}}$ or $u^h|_{\Omega_{in}}$.

---

4 for a review see [Halko, Martinsson, Tropp 11]
Approximating optimal local spaces with Randomized Linear Algebra

- Prescribe random boundary conditions; in detail choose every coefficient of a FEM basis function on $\Gamma_{out}$ as a (mutually independent) Gaussian random variable with zero mean and variance one.
- Solve PDE for random boundary conditions numerically and store evaluation of local solution of PDE $u^h|_{\Gamma_{in}}$ or $u^h|_{\Omega_{in}}$.
- Define reduced space $R_{rand}^n$ as the span of $n$ such evaluations $u^h|_{\Gamma_{in}}$ or $u^h|_{\Omega_{in}}$.

Questions: What is the quality of such an approximation? (How) can we determine the dimension of the reduced space for a given tolerance?

---

4 for a review see [Halko, Martinsson, Tropp 11]
Proposition (A priori error bound (Buhr, KS 2018))

Under the above assumptions there holds for $n, p \geq 2$

$$
\mathbb{E} \left[ \sup_{\xi \in S^h} \inf_{\zeta \in R_{\text{rand}}^{n+p}} \frac{\| T^h \xi - \zeta \|_R}{\| \xi \|_S} \right] \leq C_h \left\{ \left( 1 + \frac{\sqrt{n}}{\sqrt{p-1}} \right) \sqrt{\lambda_{n+1}^h} + \frac{e^{\sqrt{n+p}}}{p} \left( \sum_{j>n} \lambda_j^h \right)^{1/2} \right\}
$$

$$\sim c \sqrt{n} \sqrt{\lambda_{n+1}^h}$$

Optimal convergence rate achieved with transfer eigenvalue problem:

$$d_n(T(S); R) = \sup_{\xi \in S} \inf_{\zeta \in R^n} \frac{\| T \xi - \zeta \|_R}{\| \xi \|_S} = \sqrt{\lambda_{n+1}}$$

\(^5\text{based on results in [Halko, Martinsson, Tropp 11]}\)
Proposition (A priori error bound (Buhr, KS 2018))

Under the above assumptions there holds for $n, p \geq 2$

$$
\mathbb{E} \left[ \sup_{\xi \in S^h} \inf_{\zeta \in R_{\text{rand}}^{n+p}} \frac{\| T^h \xi - \zeta \|_R}{\| \xi \|_S} \right] \leq C_h \left\{ \left( 1 + \frac{\sqrt{n}}{\sqrt{p-1}} \right) \sqrt{\lambda_{n+1}^h} + \frac{e \sqrt{n+p}}{p} \left( \sum_{j>n} \lambda_j^h \right)^{1/2} \right\}
$$

\[ \sim c \sqrt{n} \sqrt{\lambda_{n+1}^h} \]

where

- $C_h = \sqrt{\frac{\lambda_{\text{max}}(M_R)}{\lambda_{\text{min}}(M_R)}} \sqrt{\frac{\lambda_{\text{max}}(M_S)}{\lambda_{\text{min}}(M_S)}}$
- $(M_R)_{i,j} = (\psi_j, \psi_i)_R, \psi_i: \text{FE basis functions}$
- $(M_S)_{i,j} = (\psi_j, \psi_i)_S, \psi_i: \text{FE basis functions}$
- $p$: oversampling parameter
Probabilistic a posteriori error bound\textsuperscript{6}

Proposition (Probabilistic a posteriori error bound (Buhr, KS 2018))

- $\{\omega^{(i)} : i = 1, 2, \ldots, n_t\}$: standard Gaussian vectors
- $D_S : \mathbb{R}^{N_{out}} \rightarrow S^h; (c_1, \ldots, c_{N_{out}}) \mapsto \chi, \chi = \sum_{i=1}^{N_{out}} c_i \psi_i, \psi_i : FE$ basis functions

Define

$$\Delta(n_t, \delta_{tf}) := \frac{c_{\text{est}}(n_t, \delta_{tf})}{\sqrt{\lambda_{\text{min}}^{M_S}}} \max_{i=1,\ldots,n_t} \left( \inf_{\zeta \in R^{n}_{\text{rand}}} \| T^{h} D_S \omega^{(i)} - \zeta \|_R \right)$$

Then there holds

$$\sup_{\xi \in S^h} \inf_{\zeta \in R^{n}_{\text{rand}}} \frac{\| T^{h} \xi - \zeta \|_R}{\| \xi \|_S} \leq \Delta(n_t, \delta_{tf}) \leq \left( \frac{\lambda_{\text{max}}^{M_S}}{\lambda_{\text{min}}^{M_S}} \right)^{1/2} c_{\text{eff}}(n_t, \delta_{tf}) \sup_{\xi \in S^h} \inf_{\zeta \in R^{n}_{\text{rand}}} \frac{\| T^{h} \xi - \zeta \|_R}{\| \xi \|_S}$$

with a probability of at least $1 - \delta_{tf}$.

\textsuperscript{6}Estimator extends results in [Halko, Martinsson, Tropp 11]; effectivity bound new
Adaptive randomized range finder

- **Input:** Select tolerance $tol$, failure probability $\delta_{\text{algofail}}$
- **While** $\Delta(n_t, \delta_{tf}) > tol$
  - Generate random boundary values on $\Gamma_{out}$
  - Apply transfer operator $T^h$ to random boundary conditions
  - Add new solution to $R^n_{\text{rand}}$
  - Orthonormalize solutions
  - Update a posteriori error estimator
- **Output:** $R^n_{\text{rand}}$ such that $\sup_{\xi \in S^h} \inf_{\zeta \in R^n_{\text{rand}}} \frac{\|T^h\xi - \zeta\|_R}{\|\xi\|_S} \leq tol$ with probability at least $1 - \delta_{\text{algofail}}$

adapted from [Halko, Martinsson, Tropp 11]
Numerical Experiments for analytic test problem

Numerical Experiments: interfaces

- Local (oversampling) domain $\Omega := (-1, 1) \times (0, 1)$
- Consider PDE: $-\Delta u = 0$ in $\Omega$
- Goal: Construct reduced space on $\Gamma_{in}$

![Diagram of $\Omega$ with $\Gamma_{in}$ and $\Gamma_{out}$]
Heat conduction: \(-\Delta u = 0\) on \(\Omega = (-1, 1) \times (0, 1)\)

**Figure:**
- Optimal basis
- Basis generated by randomized range finder algorithm
Heat conduction: \(-\Delta u = 0\) on \(\Omega = (-1, 1) \times (0, 1)\)
Heat conduction: \(-\Delta u = 0\) on \(\Omega = (-1, 1) \times (0, 8)\)

**CPU times**

### Properties of basis generation

<table>
<thead>
<tr>
<th></th>
<th>Algorithm 2</th>
<th>Scipy/ARPACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>(resulting) basis size (n)</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>operator evaluations</td>
<td>59</td>
<td>79</td>
</tr>
<tr>
<td>adjoint operator evaluations</td>
<td>0</td>
<td>79</td>
</tr>
<tr>
<td>execution time in s (without factorization)</td>
<td>20.4 s</td>
<td>47.9 s</td>
</tr>
</tbody>
</table>

**Table:** CPU times; Target accuracy \(\text{tol} = 10^{-4}\), number of testvectors \(n_t = 20\), failure probability \(\delta_{\text{algofail}} = 10^{-15}\); unknowns of corresponding problem 638,799
Numerical Experiments for a transfer operator with slowly decaying singular values

Numerical Experiments: subdomains

- local (oversampling) domain $\Omega := (-2, 2) \times (-0.25, 0.25) \times (-2, 2)$
- Consider PDE: linear elasticity in $\Omega$ (isotropic, homogeneous)
- Goal: Construct reduced space on $\Omega_{in} = (-0.5, 0.5) \times (-0.25, 0.25) \times (-0.5, 0.5)$

Figure: $\Omega \setminus \Omega_{in}$
Linear elasticity on $\Omega := (-2, 2) \times (-0.5, 0.5) \times (-2, 2)$

**Figure:** Convergence behavior of adaptive algorithm (left) and effectivity of a posteriori error estimator $\Delta/\| T^h - P_{R_{rand}}^n T^h \|$ (right) for increasing number of test vectors $n_t$. 
Olimex A64: Maxwell’s equation (results by Andreas Buhr)

- global discretization: about 65 million degrees of freedom
- 1120 subdomains
Error Estimator Decay

\[
\max_{i \in \{1, \ldots, n_t\}} \left\| (T^h - P_{R_{\text{rand}}^n}^h) \omega_i \right\|_R
\]

\[
\begin{array}{c}
0 \\
10^{-6}
\end{array}
\begin{array}{c}
10^0 \\
10^2
\end{array}
\begin{array}{c}
10^{-2}
\end{array}
\begin{array}{c}
10^{-4}
\end{array}
\begin{array}{c}
10^{-6}
\end{array}
\]

\[
\begin{array}{c}
0 \\
50 \\
100 \\
150 \\
200
\end{array}
\]
CPU timings

(On laptop)

-assembly
-factorization
-testvector generation
-basis generation

Domain vs. time/s
Outline

- Projection-based model order reduction in a nutshell
  - Randomized error estimation
- Localized Model Order Reduction
  - Constructing optimal local approximation spaces (in space)
  - Approximating optimal local approximation spaces via random sampling
  - **Generating quasi-optimal local approximation spaces in time by random sampling**

References:

Decay behavior of solutions of certain PDEs in time

- The solution space of certain system of ordinary/partial differential equations in time is locally low-rank
  - Consider

\[
\partial_t u - \text{div}(\kappa(x, t) \nabla u) = 0, \quad \text{in } D \times (0, T), \\
u(x, t) = 0 \text{ on } \partial D, \quad u(x, 0) = u_0(x).
\]

- There holds: \( \| u(\cdot, t) \|_{L^2(D)} \leq e^{-C(\kappa)t} \| u_0 \|_{L^2(D)}. \)

- **Idea:** Exploit decay behavior to efficiently construct local reduced or multiscale spaces in time.
A compact transfer operator for time-dependent problems

- Define transfer operator $T_{0 \to t^*} : L^2(D) \to \mathcal{H}_{t^*}$ that solves PDE for arbitrary initial conditions and evaluates corresponding solution in $t^*$, where

$$\mathcal{H}_{t^*} := \{ w(\cdot, t^*) \in L^2(D) : w \text{ solves PDE with } w(\cdot, 0) \in L^2(D), f \equiv 0 \}.$$
A compact transfer operator for time-dependent problems

- Define transfer operator \( T_{0\to t^*} : L^2(D) \to \mathcal{H}_{t^*} \) that solves PDE for arbitrary initial conditions and evaluates corresponding solution in \( t^* \), where

  \[
  \mathcal{H}_{t^*} := \{ w(\cdot, t^*) \in L^2(D) : w \text{ solves PDE with } w(\cdot, 0) \in L^2(D), f \equiv 0 \}.
  \]

- Heat equation with rough coefficients: \( T_{0\to t^*} \) is compact thanks to the Caccioppoli inequality:

Proposition (Caccioppoli inequality in time (KS, terMaat 2020))

Let \( w \) satisfy the weak form of the heat equation with right-hand side \( f \equiv 0 \) and arbitrary initial conditions \( w(x, 0) \) and let \( \varrho \in \mathbb{R} \) with \( \varrho > 0 \). Then, we have

\[
\| w(\cdot, t^*) \|_{L^2(D)}^2 + \| \kappa^{1/2} \nabla w \|_{L^2((\varrho, T-\varrho), L^2(D))} \leq \frac{1}{\varrho} \| w \|_{L^2(I, L^2(D))}^2.
\]
A compact transfer operator for time-dependent problems

- Define transfer operator $T_{0 \to t^*} : L^2(D) \to \mathcal{H}_{t^*}$ that solves PDE for arbitrary initial conditions and evaluates corresponding solution in $t^*$, where

$$
\mathcal{H}_{t^*} := \{ w(\cdot, t^*) \in L^2(D) : w \text{ solves PDE with } w(\cdot, 0) \in L^2(D), f \equiv 0 \}.
$$

- Heat equation with rough coefficients: $T_{0 \to t^*}$ is compact thanks to the Caccioppoli inequality.

**Proposition (Optimal approximation spaces (KS, terMaat 2020))**

The optimal approximation space in $\mathcal{H}_t$ is given by

$$
\mathcal{H}_{t^*}^n := \text{span}\{ \phi_{1}^{t^*}, \ldots, \phi_{n}^{t^*} \}, \quad \text{where } \phi_{j}^{t^*} = T_{0 \to t^*} \varphi_{j}^{t^*}, \quad j = 1, \ldots, n,
$$

and $\varphi_{j}^{t^*}$ eigenfunctions of the transfer eigenvalue problem: Find $(\varphi_{j}^{t^*}, \lambda_{j}^{t^*}) \in (\mathcal{H}_0, \mathbb{R}^+) \text{ such that}$

$$
( T_{0 \to t^*} \varphi_{j}^{t^*}, T_{0 \to t^*} w )_{L^2(D)} = \lambda_{j}^{t^*} ( \varphi_{j}^{t^*}, w )_{L^2(D)} \quad \forall w \in \mathcal{H}_0.
$$
Approximation of optimal spaces by random sampling

- Apply $T_{0 \rightarrow t^*}$ to $n$ mutually independent random initial conditions.
- Start collecting snapshots after a certain amount of time steps to let higher frequencies decay.
- Add snapshots of simulation with prescribed initial condition $u_0$ for few time steps to snapshot set.
- Apply SVD to collection of all snapshots to construct reduced space.
Approximation via random sampling for time-dependent data

- To capture time-dependent data start at different points in time
- Define transfer operator $T_{t_i \rightarrow t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time $t_i$ and evaluates corresponding solution in $t_j$
Approximation via random sampling for time-dependent data

- To capture time-dependent data start at different points in time
- Define transfer operator $T_{t_i \rightarrow t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time $t_i$ and evaluates corresponding solution in $t_j$
- Theory for $T_{0 \rightarrow t^*}$ can directly be extended to $T_{t_i \rightarrow t_j}$
Approximation via random sampling for time-dependent data

- To capture time-dependent data start at different points in time
- Define transfer operator $T_{t_i \rightarrow t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time $t_i$ and evaluates corresponding solution in $t_j$
- Theory for $T_{0 \rightarrow t^*}$ can directly be extended to $T_{t_i \rightarrow t_j}$
- Choose $n$ random points of time $t_i, i = 1, \ldots, n$ and apply $T_{t_i \rightarrow t_j}$ to a random initial condition (mutually independent).
- Apply SVD to collection of all snapshots to construct reduced space.
Approximation via random sampling for time-dependent data

- To capture time-dependent data start at different points in time
- Define transfer operator $T_{t_i \rightarrow t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time $t_i$ and evaluates corresponding solution in $t_j$
- Theory for $T_{0 \rightarrow t^*}$ can directly be extended to $T_{t_i \rightarrow t_j}$
- Choose $n$ random points of time $t_i$, $i = 1, \ldots, n$ and apply $T_{t_i \rightarrow t_j}$ to a random initial condition (mutually independent).
- Apply SVD to collection of all snapshots to construct reduced space.
- Advantage: reduced models can be constructed in parallel
Numerical experiments: Stove problem

- $\Omega = (0, 1) \times (0, 1)$, final time $T = 10$
- Consider:

\[
\partial_t u(x, y, t) - \Delta u(x, y, t) = f(x, y, t) \quad \text{in } \Omega \times (0, T),
\]
\[
u = 0 \quad \text{on } \partial \Omega \times (0, T),
\]
\[
u(x, y, 0) = \sum_{k=2}^{4} \sin(k\pi x) \sin(k\pi y).
\]

- Use FEM with $h = 0.01$ in $x$- and $y$-direction, implicit Euler with 300 time steps
Numerical experiments: solution at different points of time
Randomized local model reduction (time)

Numerical experiments: error, singular values, random starting points in time $t_i$

- Consider 10 different random starting points
- Collecting snapshots between the 12th and 15th time step after $t_i$
  $\implies$ Dimension of reduced space is 17
Randomized error estimators build on concentration inequalities for Gaussian maps can provide
- ... a very accurate estimate of the error at high probability
- ... at low cost.

Localized model order reduction: Exploit decay behavior of solutions of certain PDEs to construct optimal local approximation spaces

Randomized methods are well suited to approximate the range of maps that are low-rank; Examples: local solution spaces in space or time
- Probabilistic a priori error bound/Numerical experiments for local solution in space: convergence rate is only slightly worse compared to the optimal rate (factor $\sqrt{n}$)
- required number of local solutions of PDE scale (roughly) with size of the reduced space; Numerical experiments: faster than Lanczos
Randomized error estimators build on concentration inequalities for Gaussian maps can provide

- a very accurate estimate of the error at high probability
- at low cost.

Localized model order reduction: Exploit decay behavior of solutions of certain PDEs to construct optimal local approximation spaces

Randomized methods are well suited to approximate the range of maps that are low-rank; Examples: local solution spaces in space or time

- Probabilistic a priori error bound/Numerical experiments for local solution in space: convergence rate is only slightly worse compared to the optimal rate (factor $\sqrt{n}$)
- required number of local solutions of PDE scale (roughly) with size of the reduced space; Numerical experiments: faster than Lanczos

Thank you very much for your attention!
## Comparison with Krylov subspace methods

<table>
<thead>
<tr>
<th></th>
<th>randomized methods</th>
<th>Krylov subspace methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>computational costs</strong></td>
<td>( T_{\text{mult}}(k + n_t) + \mathcal{O}(k^2 m) )</td>
<td>ideally ( T_{\text{mult}}(k) + \mathcal{O}(k^2 (m + n)) )</td>
</tr>
<tr>
<td>stage A: ( T_{\text{mult}}(k + n_t) + \mathcal{O}(k^2 m) )</td>
<td>stage B: ( T_{\text{mult}}(k) + \mathcal{O}(k^2 (m + n)) )</td>
<td></td>
</tr>
<tr>
<td><strong>stability</strong></td>
<td>inherently stable</td>
<td>inherently unstable</td>
</tr>
<tr>
<td><strong>parallelizable</strong></td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>
Numerical experiments: Linear elasticity with $\dim(\mathcal{P}) = 20$

- Consider $-\text{div}(E(m)C : \varepsilon(u(m))) = f$ in $\Omega$ with
  - $C$ stiffness and $\varepsilon$ strain tensor
  - vertical unitary linear forcing $f$ (red arrows)
  - zero Dirichlet boundary conditions at $\|\|$.

- $E(m)$: log-normally distributed random field on $\Omega$, use truncated Karhunen-Loève decomposition with 20 terms.

- We use a tensor-based model reduction method (PGD) and estimate the relative root mean square error.
Steering the (primal) model reduction approximation

- $\Delta^\text{res}_M$: dual norm of residual divided by dual norm of r.h.s. (no inf-sup)
- $\Delta^\text{stag}_{M,k}$: relative hierarchical error estimator using $k$ increments