(Randomized) Localized Model Order Reduction

Kathrin Smetana (University of Twente)

March 24, 2020

ICERM Workshop "Algorithms for Dimension and Complexity Reduction"

Collaborators

Andreas Buhr

(formerly University of Münster)

Anthony T Patera

(MIT)

Julia Schleuß

(University of Münster)

Lukas ter Maat

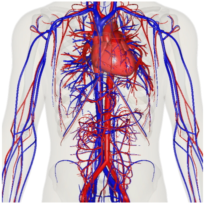
(University of Twente)

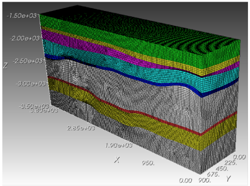
Olivier Zahm

(INRIA)

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 $\partial\Omega$
- $H_0^1(\Omega)^d \subset X \subset H^1(\Omega)^d$ $(d = 1, 2, 3)$; 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 $\mu \in \mathcal{P}$.

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.

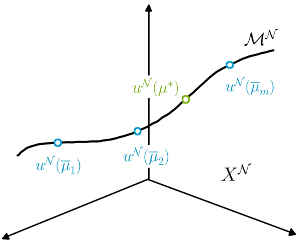
$$A(\mu)u(\mu) = f(\mu) \quad \text{in } X'.$$

- 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.

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

- Issue: Require $u^{\mathcal{N}}(\mu)$ in real time and/or for many $\mu \in \mathcal{P}$.

Projection-based model order reduction: key concept



- Exploit: $u^{\mathcal{N}}(\mu)$ belongs to "solution manifold"$\mathcal{M}^{\mathcal{N}} = \{u^{\mathcal{N}}(\mu) \,|\, \mu \in \mathcal{P}\} \subset X^{\mathcal{N}}$ of typically very low dimension
- Offline: Construct reduced space X^N from solutions $u^{\mathcal{N}}(\overline{\mu}_i)$, $i = 1, \ldots, N$ (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.

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

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 $y_j = Sr_j \in \mathbb{R}^m \quad j = 1, \dots, k$.
- Orthonormalize $y_j \longrightarrow q_j$, $j = 1, \dots, k$ and define $Q = [q_1, \dots, q_k]$
- Result: If S has exactly rank k then q_j, $j = 1, \dots, k$ span the range of S at high probability. But also in the general case q_j, $j = 1, \dots, k$ 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 of the small matrix $C = \widetilde{U}\Sigma V^*$ and set $B = Q\widetilde{U}$

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 $y_j = Sr_j \in \mathbb{R}^m \quad j = 1, \dots, k$.
- Orthonormalize $y_j \longrightarrow q_j$, $j = 1, \dots, k$ and define $Q = [q_1, \dots, q_k]$
- Result: If S has exactly rank k then q_j, $j = 1, \dots, k$ span the range of S at high probability. But also in the general case q_j, $j = 1, \dots, k$ 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 of the small matrix $C = \widetilde{U}\Sigma V^*$ and set $B = Q\widetilde{U}$

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 $\widetilde{\Delta}_N(\mu) = \beta_{LB}(\mu)^{-1}\|f(\mu) - A(\mu)u^N(\mu)\|_{X^{\mathcal{N}\prime}}$, *with* $\beta_{LB}(\mu) \leqslant \beta_{\mathcal{N}}(\mu)$ *satisfies*

$$\|u^{\mathcal{N}}(\mu) - u^N(\mu)\|_X \leqslant \widetilde{\Delta}_N(\mu) \leqslant \frac{\gamma_{\mathcal{N}}(\mu)}{\beta_{LB}(\mu)}\|u^{\mathcal{N}}(\mu) - u^N(\mu)\|_X,$$

where $\beta_{\mathcal{N}}(\mu) := \inf\limits_{v\in X^{\mathcal{N}}} \sup\limits_{w\in X^{\mathcal{N}}} \frac{\langle A(\mu)v,w\rangle}{\|v\|_X\|w\|_X}$ *and* $\gamma_{\mathcal{N}}(\mu) = \sup\limits_{v\in X^{\mathcal{N}}} \sup\limits_{w\in X^{\mathcal{N}}} \frac{\langle A(\mu)v,w\rangle}{\|v\|_X\|w\|_X}$.

- 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:

- KS, Zahm, Patera, Randomized residual-based error estimators for parametrized equations. SIAM J. Sci. Comput., 2019.
- KS, Zahm, Randomized residual-based error estimators for the proper generalized decomposition approximation of parametrized problems, Internat. J. Numer. Methods Engrg., to appear, 2020.

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

$$\mathbb{P}\left\{(1-\varepsilon)\|x-y\|_2^2 \leqslant \|\Phi x - \Phi y\|_2^2 \leqslant (1+\varepsilon)\|x-y\|_2^2 \quad \forall x, y \in \mathcal{S}\right\} \geqslant 1-\delta.$$

see for instance [Boucheron, Lugosi, Massart 2012], [Vershynin 2018]

Assumptions on random vector

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

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

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

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

Assumptions on random vector

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

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

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

$\Longrightarrow$ $(Z^{T}v)^{2}$ 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 Σ

Assumptions on random vector

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

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

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

$\Longrightarrow$ $(Z^T v)^2$ 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 Σ
- $Z_1, \ldots, Z_K$: K independent copies of Z
- Consider the following (unbiased) Monte-Carlo estimator of $\|v\|_{\Sigma}^2$

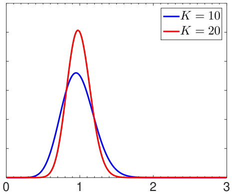
$$\frac{1}{K} \sum_{i=1}^{K} (Z_i^T v)^2.$$

Proposition (Concentration inequality (KS, Zahm, Patera 2019))

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

$$K \geqslant \frac{\log(\#\mathcal{S}) + \log(\delta^{-1})}{\log(w/\sqrt{e})} \qquad \textit{that}$$

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



- 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 $\mathcal{S} = \{\mu_1, \dots, \mu_S\} \subset \mathcal{P}$, *a failure probability* $0 < \delta < 1$, $w \in \mathbb{R}$, $w > \sqrt{e}$, *we have for*

$$K \geqslant \frac{\log(\#\mathcal{S}) + \log(\delta^{-1})}{\log(w/\sqrt{e})} \qquad \textit{that}$$

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

	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 10$
$\#\mathcal{S} = 1$	24	8	6	5	3
$\#\mathcal{S} = 100$	48	16	11	9	6
$\#\mathcal{S} = 1000$	60	20	13	11	7
$\#\mathcal{S} = 10^6$	96	31	21	17	11

Table: Values for K that guarantee (1) for all $\mu_j \in \mathcal{S}$ with $\delta = 10^{-2}$.

Proposition (Concentration inequality (KS, Zahm, Patera 2019))

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

$$K \geqslant \frac{\log(\#\mathcal{S}) + \log(\delta^{-1})}{\log(w/\sqrt{e})} \qquad \textit{that}$$

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

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

Problem: estimator $\Delta(\mu) = \left(\frac{1}{K} \sum_{i=1}^K (Z_i^T (\underline{u}^{\mathcal{N}}(\mu_j) - \underline{u}^N(\mu_j)))^2 \right)^{1/2}$
involves high-dimensional finite element solution
$\Longrightarrow$ Computationally infeasible in the online stage

A fast-to-evaluate randomized error estimator

- Exploit error residual relationship

$$Z_i^T \underline{e}(\mu) = Z_i^T \underline{A}(\mu)^{-1} \underbrace{(\underline{f}(\mu) - \underline{A}(\mu)\underline{u}^N(\mu))}_{\text{residual } \underline{r}(\mu):=} = (\underbrace{\underline{A}(\mu)^{-T} Z_i}_{\text{dual problem}})^T \underline{r}(\mu)$$

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

$$\underline{Y}_i^{\mathcal{N}}(\mu) := \underline{A}(\mu)^{-T} Z_i$$

- Approximation of the dual solutions via model order reduction:

$$\underline{y}_i^{\mathcal{N}}(\mu) \approx \underline{y}_i^{N_{du}}(\mu) \in \widetilde{\mathcal{Y}} \subset X^{\mathcal{N}}; \quad \widetilde{\mathcal{Y}}: \text{ dual reduced space.}$$

- Define fast-to-evaluate randomized error estimator

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

A fast-to-evaluate randomized error estimator

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 μ_j, $j = 1, \dots, S$ in a finite parameter set $\mathcal{S} = \{\mu_1, \dots, \mu_S\}$ and

$$K \geqslant \frac{\log(S) + \log(\delta^{-1})}{\log(w/\sqrt{e})} \qquad \textit{that} \qquad \Delta^{N_{du}}(\mu_j) := \left(\frac{1}{K} \sum_{i=1}^{K} (\underline{y}_i^{N_{du}}(\mu_j)^T \underline{r}(\mu_j))^2 \right)^{1/2}$$

satisfies

$$\mathbb{P}\Big\{ (\alpha w)^{-1} \Delta^{N_{du}}(\mu_j) \leqslant \|\underline{e}(\mu_j)\|_{\Sigma} \leqslant (\alpha w)\, \Delta^{N_{du}}(\mu_j), \quad \mu_j \in \mathcal{S}, \Big\} \geqslant 1 - \delta,$$

where

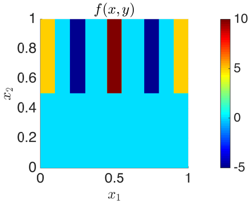
$$\alpha = \max_{\mu \in \mathcal{P}} \left(\max \left\{ \frac{\Delta(\mu)}{\Delta^{N_{du}}(\mu)}, \frac{\Delta^{N_{du}}(\mu)}{\Delta(\mu)} \right\} \right) \geqslant 1.$$

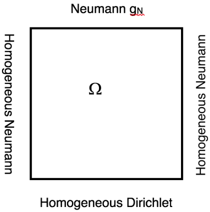
Numerical experiments: acoustics in 2D

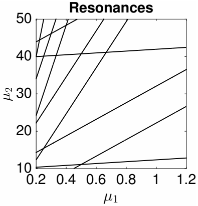
- Consider on $\Omega = (0,1) \times (0,1)$

$$
\begin{aligned}
-\partial_{x_1 x_1} u(x;\mu) - \mu_1 \partial_{x_2 x_2} u(x;\mu) - \mu_2 u(x;\mu) &= f(x) && \text{in } \Omega,\\
u(x;\mu) &= 0 && \text{on the bottom},\\
\nabla u(x;\mu) \cdot n &= 0 && \text{on the sides},\\
\kappa(\mu_1)\nabla u(x;\mu) \cdot n &= \cos(\pi x) && \text{on the top}.
\end{aligned}
$$

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

- 





- high dimensional discretization: linear FE, $h = 0.01$ in each direction

Histograms of effectivity $\Delta^{N_{du}} / \| u^{\mathcal{N}}(\mu) - u^{N}(\mu) \|_{H^1(\Omega)}$

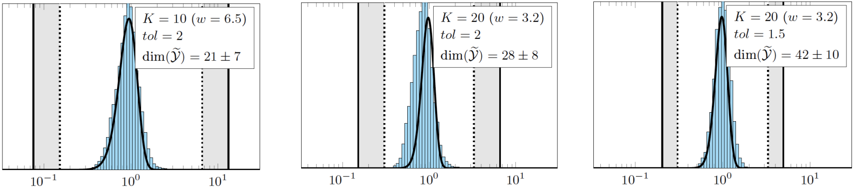


Figure: $\#\mathcal{S} = 10^4$, $N_{primal} = 20$, $q = 0.99$, 100 realizations, vertical dashed lines: $1/w$ and w, grey area: $1/(tol\, w)$ and $tol\, w$, where $\alpha \approx tol$, solid lines: chi-squared distribution

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

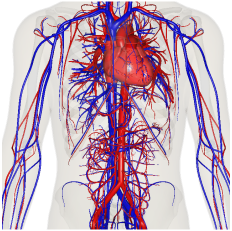
References:

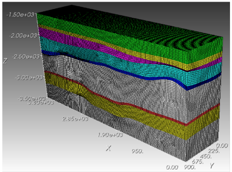
- Review: Buhr, Iapichino, Ohlberger, Rave, Schindler, and KS. Localized model reduction for parameterized problems. Invited book chapter in Handbook on Model Order Reduction. Walter De Gruyter GmbH, Berlin, 2020; also on arXiv.
- KS, Patera, Optimal local approximation spaces for component-based static condensation procedures, SIAM J. Sci. Comput., 2016.

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

$\longrightarrow$ **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)

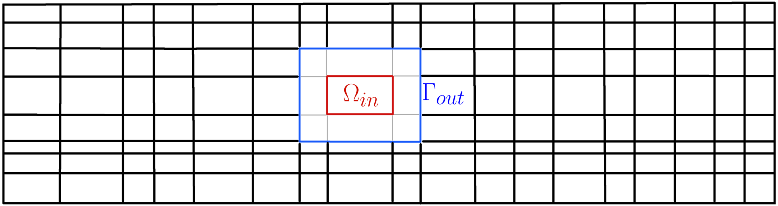
$\Longrightarrow$ Yields superalgebraic convergence and rigorous error analysis

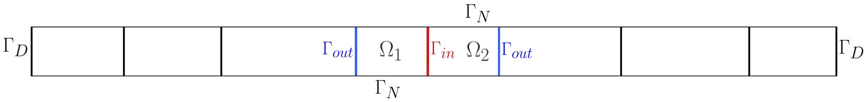
Localized model order reduction

Challenges:

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

$\Longrightarrow$ Infinite dimensional parameter space





Localized model order reduction

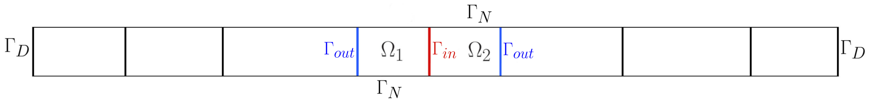
Challenges:

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

$\Longrightarrow$ Infinite dimensional parameter space

Idea:

- Restrict to space of functions that solve the PDE locally on Ω for arbitrary boundary conditions on Γ_{out}
- Exploit that for those local solutions we have a very fast decay of higher frequencies from Γ_{out} to Ω_{in}, Γ_{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 $\dim R^n \leqslant n$, 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 $\Omega = (-5,5) \times (0,1)$

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

- plus: arbitrary Dirichlet boundary conditions on Γ_{out}.

Motivation: separation of variables

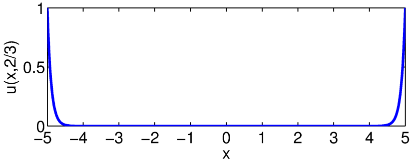
- Consider $\Omega = (-5, 5) \times (0, 1)$

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

- plus: arbitrary Dirichlet boundary conditions on Γ_{out}.
- separation of variables: all harmonic functions on Ω 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 Γ_{out} and thus $n = 3$:



Motivation: separation of variables

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

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

- plus: arbitrary Dirichlet boundary conditions on Γ_{out}.
- separation of variables: all harmonic functions on Ω have the form

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

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

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

$\Longrightarrow$ A reduced space of very low dimension on Γ_{in} will already yield a very good approximation!

Motivation: separation of variables

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

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

- plus: arbitrary Dirichlet boundary conditions on Γ_{out}.
- separation of variables: all harmonic functions on Ω have the form

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

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

Question: How can we generalize this idea?

The space of all local solutions of the PDE on Ω

- Consider the space of all local solutions of the PDE[1] on Ω

$$\mathcal{H} := \{w \in H^1(\Omega) : \text{with } \; Aw = 0 \quad \in X'\}.$$

- global solution of the PDE restricted to Ω lies in $\mathcal{H}$!
- We are interested in $u|_{\Gamma_{in}}$ or $u|_{\Omega_{in}}$ and thus introduce

$$R := \{w|_{\Gamma_{in}}, \;\; w \in \mathcal{H}\} \quad \text{or} \quad R := \{w|_{\Omega_{in}}, \;\; w \in \mathcal{H}\},$$
$$\text{and} \quad S := \{w|_{\Gamma_{out}}, \;\; w \in \mathcal{H}\}.$$

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

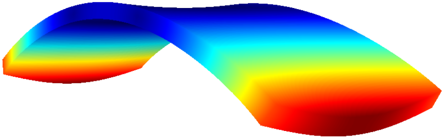
Transfer operator

- We introduce a transfer operator

$$T : S \to R$$

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

$$T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}.$$



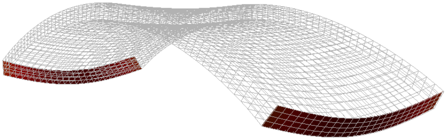
Transfer operator

- We introduce a transfer operator

$$T : S \to R$$

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

$$T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}.$$



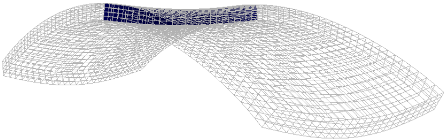
Transfer operator

- We introduce a transfer operator

$$T : S \to R$$

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

$$T(w|_{\Gamma_{out}}) := w|_{\Gamma_{in}} \quad \text{or} \quad T(w|_{\Gamma_{out}}) := w|_{\Omega_{in}}.$$



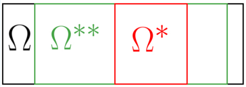
Transfer operator

- We introduce a transfer operator $T : S \to R$
- 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:

Lemma (Caccioppoli inequality for heat conduction)

Let $\kappa \in L^{\infty}(\Omega)$ *fulfill* $0 < \kappa_0 \leqslant \kappa \leqslant \kappa_1$ *with constants* κ_0, κ_1, *define* $X^0 = \{v \in H^1(\Omega), v|_{\Gamma_{out}} = 0\}$, *let* $u \in X := \{v \in H^1(\Omega), v|_{\Gamma_{out}} = g\}$ *satisfy*

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



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

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

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

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

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

$$\left(\; \varphi_j|_{D_{in}} \;\; , \;\; w|_{D_{in}} \;\; \right)_R = \lambda_j \;\left(\; \varphi_j|_{\Gamma_{out}} \;\; , \;\; w|_{\Gamma_{out}} \right)_S \;\; \forall w \in \mathcal{H}, D_{in} = \Gamma_{in}, \Omega_{in}$$

Transfer eigenvalue problem

Proposition (Transfer eigenvalue problem)

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

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

- *List λ_j such that $\lambda_1 \geqslant \lambda_2 \geqslant \ldots$, and $\lambda_j \to 0$ as $j \to \infty$.*
- *The optimal space on Γ_{in} or Ω_{in} is given by*

$$R^n := \text{span}\{\phi_1^{sp}, \ldots, \phi_n^{sp}\}, \ \phi_j^{sp} = T\varphi_j|_{\Gamma_{out}}, \quad j = 1, \ldots, n.$$

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

A priori error bound

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

u: (exact) solution,
u^n: continuous port reduced static condensation solution employing the optimal port space R^n.
We have:

$$\frac{|||u - u^n|||}{|||u|||} \leqslant C_1(\Omega)\sqrt{\lambda_{n+1}},$$

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

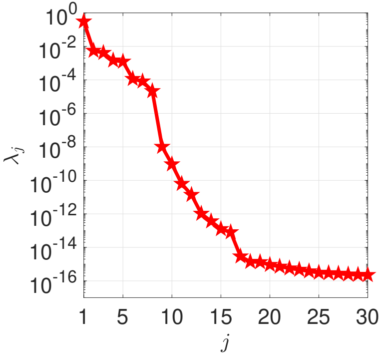


Figure: eigenvalues λ_n

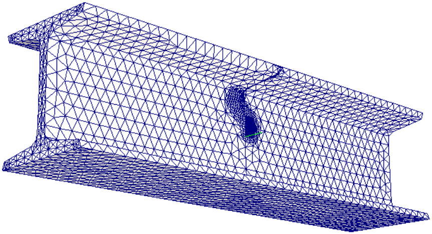


Figure: component mesh

Numerical experiments for isotropic linear elasticity

Stiffened plate — simplified model for ship stiffener

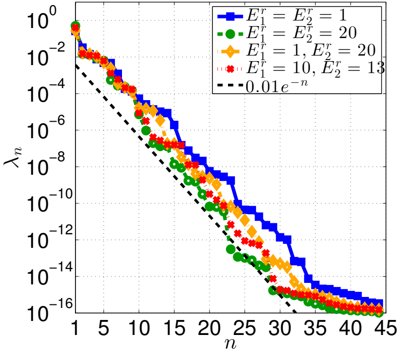


Figure: eigenvalues λ_j

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

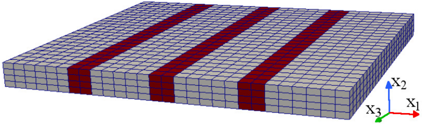


Figure: mesh in Ω_i

Numerical experiments for isotropic linear elasticity

Stiffened plate — simplified model for ship stiffener

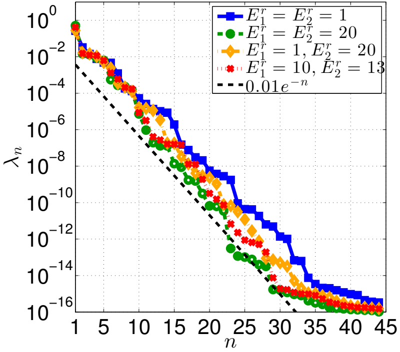


Figure: eigenvalues λ_j

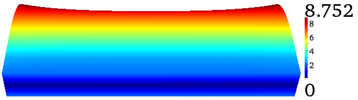


Figure: plate under bending

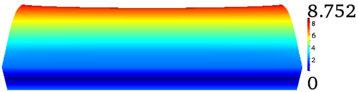


Figure: stiffened plate under bending

Comparison with other reduced interface spaces

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

- 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

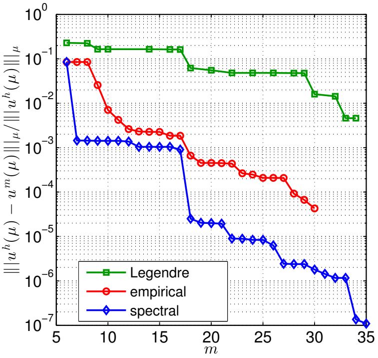
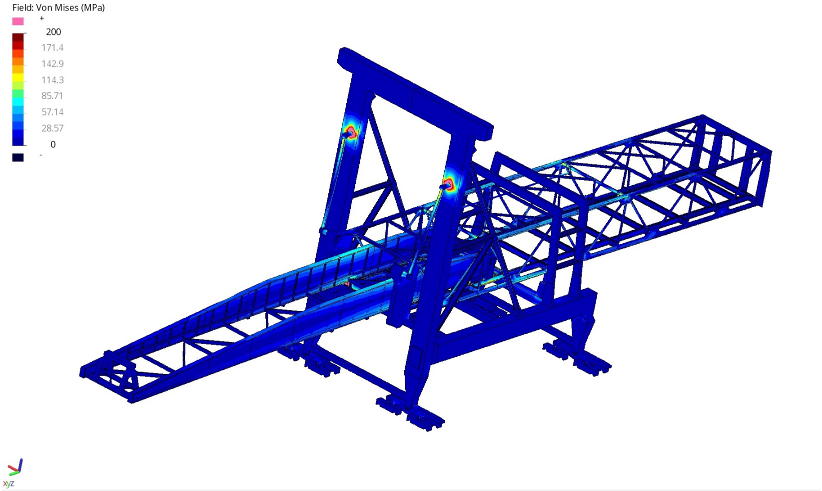


Figure: $|||u^h(\mu) - u^m(\mu)|||_\mu / |||u^h(\mu)|||_\mu$

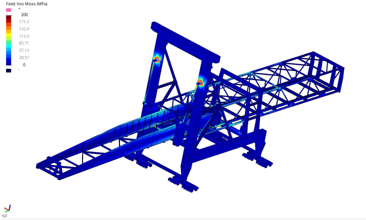
[2] $\mathcal{P}_i = [1, 10] \times [1, 1]$ for $\mu_i = (E_i, E_i^r)$

Numerical experiments: shiploader[3]



[3] Results by company Akselos S.A.; KS has no financial interest in Akselos S.A.

Numerical experiments: shiploader[3]



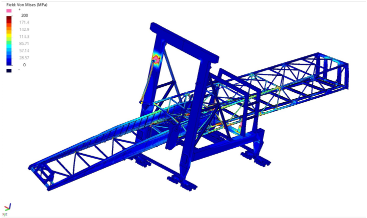


Figure: shiploader shiploader with defect

- discretization with FEM: >20 millions of DOFs
- size of Schur complement system: ≈349 000
- size of reduced Schur complement system: ≈12 000
- simulation time with reduced port spaces: ≈ 2 sec

[3] Results by company Akselos S.A.; KS has no financial interest in Akselos S.A.

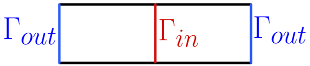
Computing an approximation of the transfer eigenvalue problem

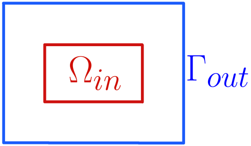
Transfer eigenvalue problem: Find $(\varphi_j, \lambda_j) \in (\mathcal{H}, \mathbb{R}^+)$ such that

$$\left(\; T^h(\varphi_j|_{\Gamma_{out}}) \;,\; T^h(w|_{\Gamma_{out}}) \; \right)_R = \lambda_j \; \left(\; \varphi_j|_{\Gamma_{out}} \;,\; w|_{\Gamma_{out}} \right)_S \;\; \forall w \in \mathcal{H}$$

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

1. Introduce a FE discretization with $\mathcal{N}_{out}$ degrees of freedom (DOFs) on Γ_{out} and $\mathcal{N}_{in}$ DOFs on Γ_{in} or Ω_{in}
2. Solve for each basis function on Γ_{out} the PDE locally
$\Longrightarrow$ number of required local solutions of the PDE scales with the number of DOFs on Γ_{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

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

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

1. Introduce a FE discretization with $\mathcal{N}_{out}$ degrees of freedom (DOFs) on Γ_{out} and $\mathcal{N}_{in}$ DOFs on Γ_{in} or Ω_{in}
2. Solve for each basis function on Γ_{out} the PDE locally
   $\Longrightarrow$ number of required local solutions of the PDE scales with the number of DOFs on Γ_{out} and thus depends on the discretization
3. Assemble and solve generalized eigenvalue problem

Problem: For large number of DOFs on Γ_{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

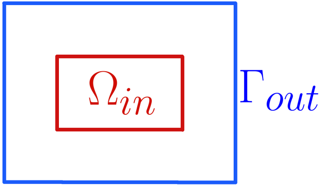
Reference: Buhr, KS, Randomized Local Model Order Reduction, SIAM J. Sci. Comput., 2018.

References on randomization in multiscale, domain decomposition methods

- Wang, Vouvakis 2015
- Calo, Efendiev, Galvis, Li 2016
- Owhadi 2015, 2017
- Chen, Li, Lu, and Wright, arXiv:1801.06938; arXiv:1807.08848

Approximating optimal local spaces with Randomized Linear Algebra[4]

- Prescribe random boundary conditions; in detail choose every coeffcient of a FEM basis function on Γ_{out} as a (mutually independent) Gaussion 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[4]

- Prescribe random boundary conditions; in detail choose every coeffcient of a FEM basis function on Γ_{out} as a (mutually independent) Gaussion 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 $u^h|_{\Omega_{in}}$

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]

Probalistic a priori error bound[5]

Proposition (A priori error bound (Buhr, KS 2018))

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

$$\mathbb{E}\left[\sup_{\xi \in S^h} \inf_{\zeta \in R^{n+p}_{rand}} \frac{\| T^h \xi - \zeta \|_R}{\| \xi \|_S}\right] \leqslant \underbrace{C_h \left\{ \left(1 + \frac{\sqrt{n}}{\sqrt{p-1}}\right) \sqrt{\lambda^h_{n+1}} + \frac{e\sqrt{n+p}}{p} \left(\sum_{j>n} \lambda^h_j\right)^{1/2} \right\}}_{\sim c\sqrt{n}\sqrt{\lambda^h_{n+1}}}$$

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] based on results in [Halko, Martinsson, Tropp 11]

Probalistic a priori error bound[5]

Proposition (A priori error bound (Buhr, KS 2018))

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

$$\mathbb{E}\left[\sup_{\xi \in S^h} \inf_{\zeta \in R^{n+p}_{rand}} \frac{\| T^h \xi - \zeta \|_R}{\|\xi\|_S}\right] \leqslant \underbrace{C_h \left\{ \left(1 + \frac{\sqrt{n}}{\sqrt{p-1}}\right) \sqrt{\lambda^h_{n+1}} + \frac{e\sqrt{n+p}}{p} \left(\sum_{j>n} \lambda^h_j\right)^{1/2} \right\}}_{\sim\, c\sqrt{n}\sqrt{\lambda^h_{n+1}}}$$

where

- $C_h = \sqrt{\frac{\lambda_{max}(\underline{M}_R)}{\lambda_{min}(\underline{M}_R)}} \sqrt{\frac{\lambda_{max}(\underline{M}_S)}{\lambda_{min}(\underline{M}_S)}}$
- $(\underline{M}_R)_{i,j} = (\psi_j, \psi_i)_R$, ψ_i: *FE basis functions*
- $(\underline{M}_S)_{i,j} = (\psi_j, \psi_i)_S$, ψ_i: *FE basis functions*
- p: *oversampling parameter*

Probablistic a posteriori error bound[6]

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

- $\{\underline{\omega}^{(i)} \ : \ i = 1, 2, ..., n_t\}$: *standard Gaussian vectors*
- $D_S : \mathbb{R}^{\mathcal{N}_{out}} \to S^h$; $(c_1, ..., c_{\mathcal{N}_{out}}) \mapsto \chi$, $\chi = \sum_{i=1}^{\mathcal{N}_{out}} c_i \psi_i$, ψ_i : *FE basis functions*

Define

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

Then there holds

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

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

[6] Estimator extends results in [Halko, Martinsson, Tropp 11]; effectivity bound new

Adaptive randomized range finder[7]

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

[7] 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 Ω
- Goal: Construct reduced space on Γ_{in}

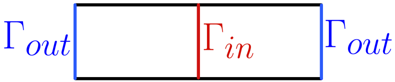


Figure: Ω

Heat conduction: $-\Delta u = 0$ on $\Omega = (-1, 1) \times (0, 1)$

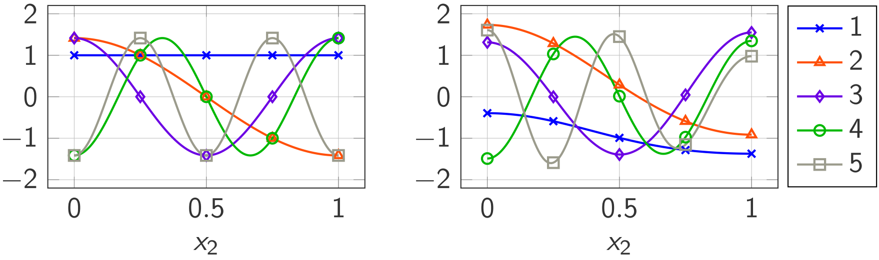
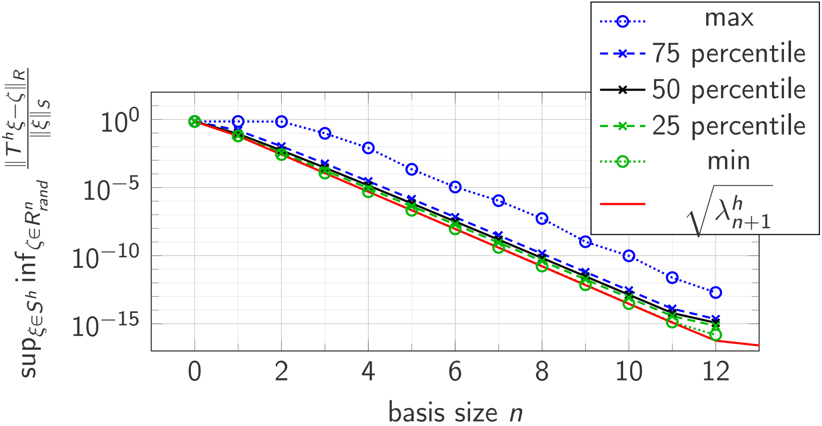


Figure: optimal basis | basis generated by randomized range finder algorithm

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

	Algorithm 2	Scipy/ARPACK
(resulting) basis size n	39	39
operator evaluations	59	79
adjoint operator evaluations	0	79
execution time in s (without factorization)	20.4 s	47.9 s

Table: CPU times; Target accuracy `tol`$= 10^{-4}$, number of testvectors $n_t = 20$, failure probability $\delta_{\mathrm{algofail}} = 10^{-15}$; 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 Ω (isotropic, homogeneous)
- Goal: Construct reduced space on $\Omega_{in} = (-0.5, 0.5) \times (-0.25, 0.25) \times (-0.5, 0.5)$

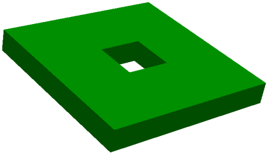


Figure: $\Omega \setminus \Omega_{in}$

Linear elasticity on $\Omega := (-2,2) \times (-0.5,0.5) \times (-2,2)$

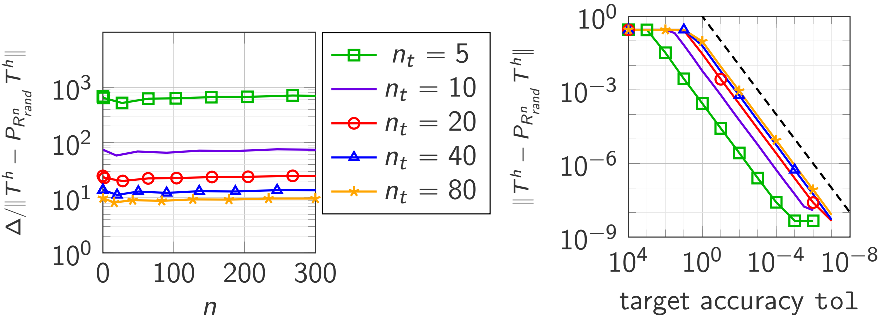
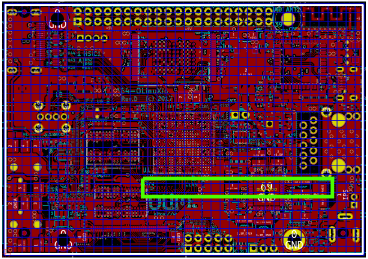


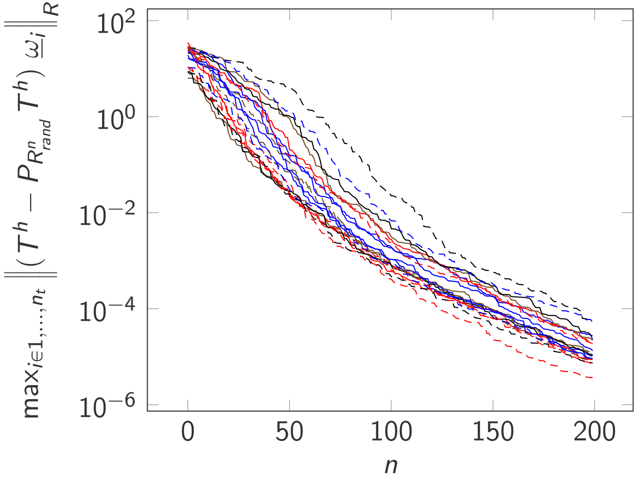
Figure: Convergence behavior of adaptive algorithm (left) and effectivity of a posteriori error estimator $\Delta / \| T^h - P_{R^n_{rand}} T^h \|$ (right) for increasing number of test vectors n_t.

Olimex A64: Maxwell's equation (results by Andreas Buhr)



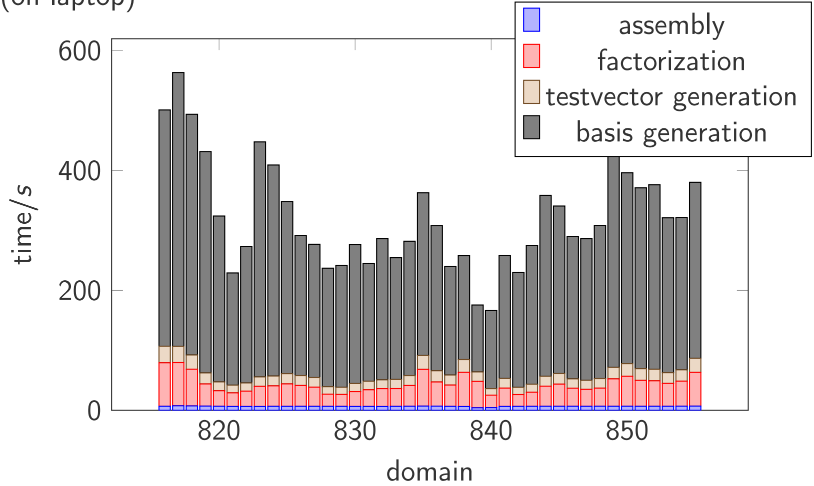
- global discretization: about 65 million degrees of freedom
- 1120 subdomains

Error Estimator Decay



CPU timings

(on laptop)



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:

- KS, Schleuß, Optimal local approximation spaces for parabolic problems, in preparation.
- KS, ter Maat, Generating quasi-optimal local approximation spaces in time by random sampling, in preparation.

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

$$\begin{aligned} \partial_t u - \operatorname{div}(\kappa(x,t)\nabla u) = 0, \quad \text{in } D \times (0,T), \\ u(x,t) = 0 \text{ on } \partial D, \quad u(x,0) = u_0(x). \end{aligned}$$

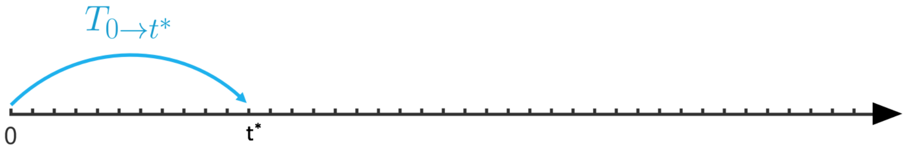
  - There holds: $\|u(\cdot,t)\|_{L^2(D)} \leqslant e^{-C(\kappa)t}\|u_0\|_{L^2(D)}$.

- 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

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



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

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

- 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

$$\|w(\cdot, t^*)\|^2_{L^2(D)} + \|\kappa^{1/2}\nabla w\|_{L^2((\varrho, T-\varrho), L^2(D))} \leqslant \frac{1}{\varrho}\|w\|^2_{L^2(I, L^2(D))}.$$

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

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

- 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

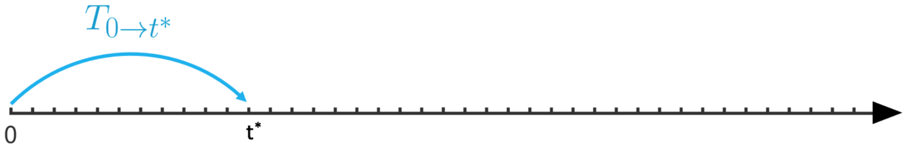
$$\mathcal{H}^n_{t*} := \mathrm{span}\{\phi_1^{t^*}, ..., \phi_n^{t^*}\}, \qquad \textit{where } \phi_j^{t^*} = \mathcal{T}_{0\to t*}\varphi_j^{t^*}, \quad j = 1, ..., n,$$

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

$$(\, \mathcal{T}_{0\to t*}\varphi_j^{t^*} \,,\, \mathcal{T}_{0\to t*} w \,)_{L^2(D)} = \lambda_j^{t^*} (\, \varphi_j^{t^*} \,,\, w \,)_{L^2(D)} \quad \forall w \in \mathcal{H}_0.$$

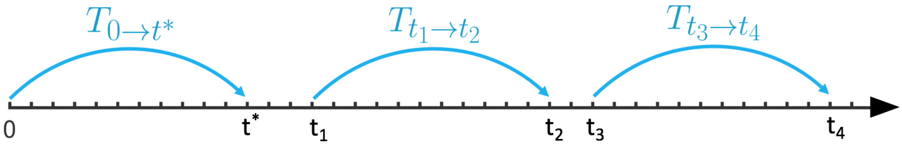
Approximation of optimal spaces by random sampling

- Apply $T_{0\to t*}$ to n mututally 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.



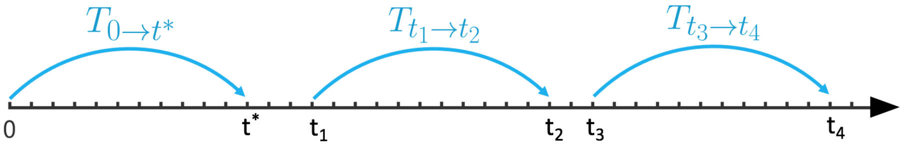
Approximation via random sampling for time-dependent data

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



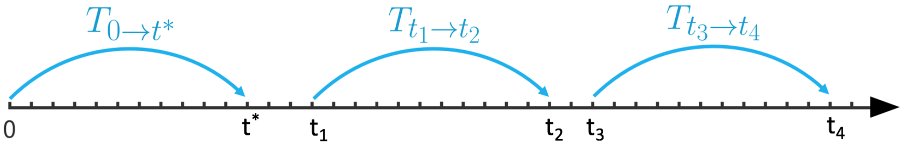
Approximation via random sampling for time-dependent data

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



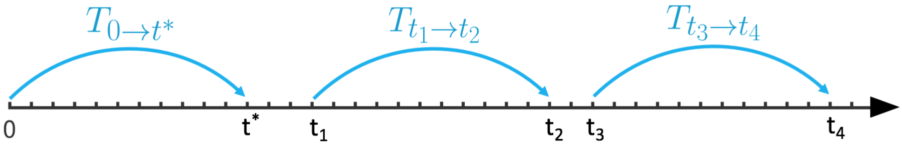
Approximation via random sampling for time-dependent data

- To capture time-dependent data start at different points in time
- Define transfer operator $T_{t_i \to t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time t_i and evaluates corresponding solution in t_j
- Theory for $T_{0 \to t^*}$ can directly be extended to $T_{t_i \to t_j}$
- Choose n random points of time t_i, $i = 1, \ldots, n$ and apply $T_{t_i \to t_j}$ 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 $T_{t_i \to t_j}$ that solves PDE for arbitrary initial conditions, arbitrary starting time t_i and evaluates corresponding solution in t_j
- Theory for $T_{0 \to t*}$ can directly be extended to $T_{t_i \to t_j}$
- Choose n random points of time t_i, $i = 1, \ldots, n$ and apply $T_{t_i \to t_j}$ 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

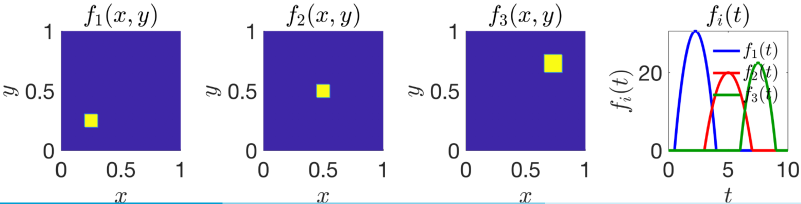


Numerical experiments: Stove problem

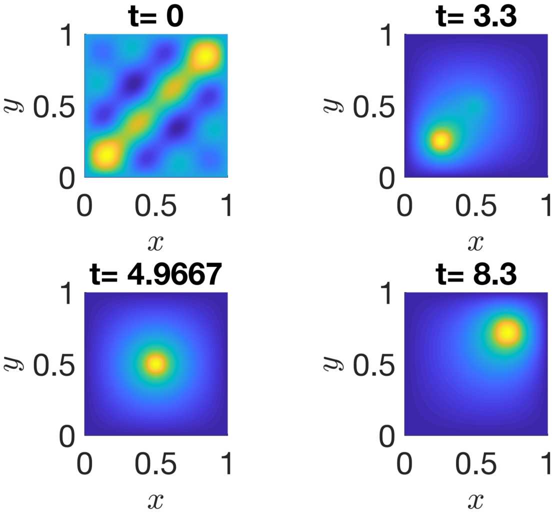
- $\Omega = (0,1) \times (0,1)$, final time $T = 10$
- Consider:

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

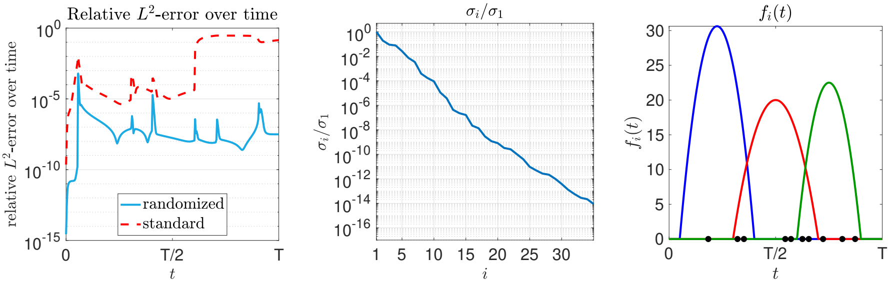
- Use FEM with $h = 0.01$ in x- and y-direction, implicit Euler with 300 time steps



Numerical experiments: solution at different points of time



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

$\Longrightarrow$ Dimension of reduced space is 17

Summary

- 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

Summary

- 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

	randomized methods	Krylov subspace methods
computational costs	stage A: $T_{mult}(k+n_t) + \mathcal{O}(k^2 m)$ stage B: $T_{mult}(k) + \mathcal{O}(k^2(m+n))$	ideally $T_{mult}(k) + \mathcal{O}(k^2(m+n))$
stability	inherently stable	inherently unstable
parallelizable	yes	no

Numerical experiments: Linear elasticity with $\dim(\mathcal{P}) = 20$

- Consider $-\mathrm{div}(E(m)C : \varepsilon(u(m))) = f$ in Ω with
  - C stiffness and ε strain tensor
  - vertical unitary linear forcing f (red arrows)
  - zero Dirichlet boundary conditions at $||||$
- $E(m)$: log-normally distributed random field on Ω, 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

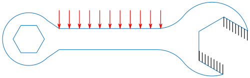


Figure: Boundary conditions, f

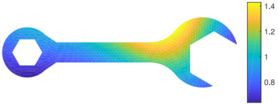


Figure: Realization of field $\log(E(m))$

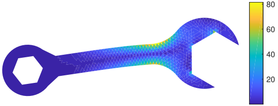
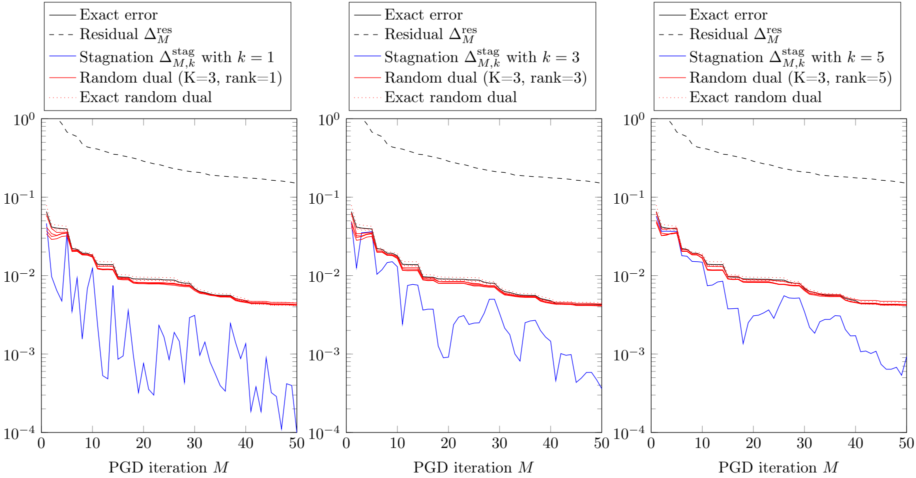


Figure: Corresponding solution

Steering the (primal) model reduction approximation



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