Randomized Model Order Reduction

Kathrin Smetana (University of Twente)
joint work with
A. Buhr (University of Münster), A. T. Patera (MIT),
and O. Zahm (INRIA)

June 8, 2018

4TU-AMI Symposium “Reducing dimensions in Big Data: Model Order Reduction in action”
Data and Model Order Reduction

- Reduced order models
  - Generate (synthetic)
  - Combine
  - Help to construct

Data (analytics)

Methods from

K Smetana (k.smetana@utwente.nl)
Motivation/Outline

- **Motivation:** Randomized methods have got a steadily growing deal of attention in recent years, especially for problems in large-scale data analysis.
  Two most important benefits:
  - They can result in faster algorithms, either in worst-case asymptotic theory and/or numerical implementation,
  - they allow very often for (novel) tight error estimators

- **Topic of this talk:** Show how we can benefit from randomized methods in model order reduction
Motivation/Outline

- **Motivation**: Randomized methods have got a steadily growing deal of attention in recent years, especially for problems in large-scale data analysis.
  
  Two most important benefits:
  - They can result in faster algorithms, either in worst-case asymptotic theory and/or numerical implementation,
  - they allow very often for (novel) tight error estimators

- **Topic of this talk**: Show how we can benefit from randomized methods in model order reduction

- **Outline**:
  1. Introduction to projection-based model order reduction
  2. Construct reduced spaces via randomized methods
  3. Randomized a posteriori error estimator for projection-based model reduction
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 $\mu \in \mathcal{P}$. 
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:** \( u^N(\mu) \) belongs to "solution manifold" \( \mathcal{M}^N = \{ u^N(\mu) | \mu \in \mathcal{P} \} \subset X^N \) of typically very low dimension

- **Offline:** Construct reduced space \( X^N \subset X^N \) 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.
  \[
  A(\mu^*) u^N(\mu^*) = f(\mu^*) \quad \text{in } X^{N'}.
  \]
  
  If \( \mathcal{M}^N \) is smooth, \( N \ll N \) already yields a very accurate approximation. ([DeVore Petrova Wojtaszczyk 13])

For an overview on model order reduction see for instance [Benner, Cohen, Ohlberger, Willcox 2017].
Proper Orthogonal Decomposition via SVD

- Introduce finite dimensional training set $P_{\text{train}} \subset P$ of dimension $n_{\text{train}}$
- Compute solutions $u^N(\mu)$ for all $\mu \in P_{\text{train}}$
- Store coefficients in a so-called snapshot matrix:
  \[ Y = \begin{bmatrix} u^N(\mu_1) & \cdots & u^N(\mu_{n_{\text{train}}}) \end{bmatrix} \]
- Perform Singular Value decomposition: $Y = U\Sigma V^*$
- Use first $N$ left singular vectors to define reduced space
Algorithm 1: Greedy algorithm

**input**: finite dimensional training set $P^{\mbox{train}} \subset P$, tolerance $tol$

**output**: $S_N$, $X^N$

**Initialize**: $S_1 = \emptyset$, $X^0 = \{0\}$, $\Delta_0(\mu) = \|u^N(\mu)\|_X$

**for** $N = 1 : N_{\mbox{max}}$ **do**

**Find**: $\mu_N = \arg \max_{\mu \in P^{\mbox{train}}} \Delta_{N-1}(\mu)$. ($\|u^N(\mu) - u^N(\mu)\|_X \leq \Delta_N(\mu)$)

Solve for $u^N(\mu_N)$.

**Extend**: $S_N = S_{N-1} \cup \mu_N$ and $X^N = \mbox{span}\{u^N(\mu_1), \ldots, u^N(\mu_N)\}$.

Compute $\Delta_N(\mu)$ for all $\mu \in P^{\mbox{train}}$.

**if** $\arg \max_{\mu \in P^{\mbox{train}}} \Delta_N(\mu) \leq tol$ **then**

**break**

**end**

**end**
Certification via a posteriori error bound

- A posteriori error estimator is important both
  - to construct reduced order models via the greedy
  - certify the approximation in the online stage: how large is the error (in some quantity of interest)?

**Proposition (A posteriori error bound)**

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

$$ \| u^N(\mu) - u^N(\mu) \|_{\mathcal{X}} \leq \Delta_N(\mu) \leq \frac{\gamma_N(\mu)}{\beta_{LB}(\mu)} \| u^N(\mu) - u^N(\mu) \|_{\mathcal{X}}, $$

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

- Problem: Good estimate of inf-sup often computationally infeasible
  $\rightarrow$ constant-free randomized error estimators
Constructing reduced order models via methods from randomized numerical linear algebra

For an overview on algorithms in randomized numerical linear algebra see for instance: [Halko et al 2011], [Mahoney 2011], [Drineas, Mahoney, 2016]
Randomized SVD

- **Goal**: Given a snapshot matrix \( Y \in \mathbb{R}^{N \times n_{\text{train}}} \) and an integer \( k \) find an orthonormal matrix \( Q \) of rank \( k \) such that \( Y \approx QQ^* Y \).

- **Approach**:
  - Draw \( k \) random vectors \( r_j \in \mathbb{R}^{n_{\text{train}}} \) (say standard Gaussian)
  - Form sample vectors \( s_j = Yr_j \in \mathbb{R}^N \quad j = 1, \ldots, k \).
  - Orthonormalize \( s_j \rightarrow q_j, = 1, \ldots, k \) and define \( Q = [q_1, \ldots, q_k] \)

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

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

References in MOR: Hochman et al 2014, Alla, Kutz 2015, Balabanov, Nouy 2018
Reduced order modelling for large-scale problems

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**
Reduced order modelling for large-scale problems

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**
Domain decomposition and oversampling

For each small subdomain $\omega_i$ (e.g. $\omega_{816}$, yellow) we introduce an oversampling domain $\omega_i^*$ (e.g. $\omega_{816}^*$, green)
Domain decomposition and oversampling

For each small subdomain $\omega_i$ (e.g. $\omega_{816}$, yellow) we introduce an oversampling domain $\omega_i^*$ (e.g. $\omega_{816}^*$, green)
Domain decomposition and oversampling

For each small subdomain $\omega_i$ (e.g. $\omega_{816}$, yellow) we introduce an oversampling domain $\omega_i^*$ (e.g. $\omega_{816}^*$, green)
Construct reduced order models via randomized methods

Localized model order reduction

Constructing local reduced models via a transfer operator

Introduce transfer operators $T_i$:

- ... acts on the space of local solutions of the PDE and maps values $\zeta$ on $\partial \omega_i^*$ to $\omega_i$

- ... by solving the PDE locally with Dirichlet boundary values $\zeta$

- ... and restricting the local solution to $\omega_i$
Constructing optimal local spaces via transfer operators

- Key observation: Global solution $u$ satisfies $u|_{\omega_i} = T_i(u|_{\partial\omega_i^*})$

$\implies$ Construct local reduced spaces that approximate $\text{range}(T_i)$

- $\phi_j$: the left singular vectors of $T_i$; $\sigma_j$: singular values of $T_i$
- The reduced space

$$R_{i,\text{opt},n} := \text{span}\{\phi_1, \ldots, \phi_n\}$$

is the optimal space and minimizes the approximation error among all spaces of the same dimension

- The error satisfies:

$$\|T_i - P_{R_{i,\text{opt},n}} T_i\| = \sigma_{n+1}, \quad P_{R_{i,\text{opt},n}} : \text{orthogonal projection on } R_{i,\text{opt},n}$$

References: Babuska, Lipton, MMS, 2011; Smetana, Patera, SISC, 2016
Approximating optimal local spaces with Randomized Linear Algebra

- Prescribe random boundary conditions; in detail choose every coefficient of a FEM basis function on $\partial \omega_i^*$ 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 $u|_{\omega_i}$.
- Define reduced space $R_{i,\text{rand}}^n$ as the span of $n$ such evaluations $u|_{\omega_i}$.

References: Buhr, Smetana, SISC, 2018
Approximating optimal local spaces with Randomized Linear Algebra

- Prescribe random boundary conditions; in detail choose every coefficient of a FEM basis function on $\partial \omega_i^*$ 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 $u|_{\omega_i}$.
- Define reduced space $R_{i,\text{rand}}^n$ as the span of $n$ such evaluations $u|_{\omega_i}$.

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

References: Buhr, Smetana, SISC, 2018
Construct reduced order models via randomized methods

Probabilistic a priori error bound

Proposition (A priori error bound (Buhr, Smetana 18))

\( T : S \rightarrow R \) transfer operator as above, \( p \) oversampling parameter, \( n, p \geq 2 \)

\[
\mathbb{E} \| T_i - P_{R_i,n+p} T_i \| \leq \sqrt{\frac{M_R}{\lambda_{\min} M_S}} \frac{M_S}{\lambda_{\max} M_R} \left\{ \left( 1 + \frac{\sqrt{n}}{\sqrt{p-1}} \right) \sigma_{n+1} + \frac{e\sqrt{n+p}}{p} \left( \sum_{j>n} \sigma_j^2 \right)^{1/2} \right\}
\]

\[ \sim c \sqrt{n} \sigma_{n+1} \]

Optimal convergence rate achieved via SVD:

\[
\| T_i - P_{R_i,\text{opt},n} T_i \| = \sigma_{n+1}
\]

\(^1\) based on results in [Halko, Martinsson, Tropp 11]
Proposition (Probabilistic a posteriori error bound (Buhr, Smetana 2018))

\[ \{ \mathbf{r}^{(j)} : j = 1, 2, ..., n_t \} : \text{standard Gaussian vectors} \]

Define

\[ \Delta(n_t, \delta_{tf}) := \frac{c_{\text{est}}(n_t, \delta_{tf})}{\sqrt{\lambda_{\min}^{MS}}} \max_{j \in 1, ..., n_t} \left( \| T_i \mathbf{r}^{(j)} - P_{R_i, \text{rand}} T_i \mathbf{r}^{(j)} \| \right) \]

Then there holds

\[ \| T_i - P_{R_i, \text{rand}} T_i \| \leq \Delta(n_t, \delta_{tf}) \leq \left( \frac{\lambda^{MS}_{\max}}{\lambda^{MS}_{\min}} \right)^{1/2} c_{\text{eff}}(n_t, \delta_{tf}) \| T_i - P_{R_i, \text{rand}} T_i \| \]

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

\[ ^2 \text{Estimator extends results in [Halko, Martinsson, Tropp 11]; effectivity bound new} \]
Adaptive randomized range finder

- **Input:** Select tolerance $tol$, failure probability $\delta_{algofail}$
- **While** $\Delta(n_t, \delta_{tf}) > tol$
  - Generate random boundary values on $\partial \omega_i^*$
  - Apply transfer operator $T_i$ to random boundary conditions
  - Add new solution to $R_{i, rand}^n$
  - Orthonormalize solutions
  - Update a posteriori error estimator
- **Output:** $R_{i, rand}^n$ such that $\| T_i - P_{R_{i, rand}^n} T_i \| \leq tol$ with probability at least $1 - \delta_{algofail}$
Construct reduced order models via randomized methods

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 interface \(\Gamma_{in}\)

Figure: \(\omega^*\)
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 $\omega^* = (-1, 1) \times (0, 1)$
Construct reduced order models via randomized methods

**Numerical experiments**

Heat conduction: \(-\Delta u = 0\) on \(\omega^* = (-1, 1) \times (0, 8)\)

**CPU times**

**Properties of basis generation**

<table>
<thead>
<tr>
<th></th>
<th>randomized</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 \( \omega = (-0.5, 0.5) \times (-0.25, 0.25) \times (-0.5, 0.5) \)

Figure: \( \omega^* \setminus \omega \)
Construct reduced order models via randomized methods

Numerical experiments

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

![Graphs showing convergence behavior of adaptive algorithm and effectivity of posteriori error estimator.]

**Figure:** Convergence behavior of adaptive algorithm (left) and effectivity of a posteriori error estimator $\Delta/\| T - P^n_{R_{rand}} T \|$ (right) for increasing number of test vectors $n_t$. 
Olimex A64

- 1.2 GHz quad-core ARM CPU
- 1 GB of RAM
- open hardware
- designed with KiCAD

Results by Andreas Buhr
Decay of Singular Values for $T_{816}$

- $\sigma_{n+1}$
- $\sigma_{n+1}\sqrt{n}$
- $\Delta$

$n$ values range from 0 to 200.
Construct reduced order models via randomized methods

Error Estimator Decay

K Smetana (k.smetana@utwente.nl)
Error Estimator Decay

\[ \max_{i=1, \ldots, n_t} \| (T - P_{R_{\text{rand}}}^n) r_i \|_R \]

\[ \log_{10}(n) \]

\[ \log_{10}(\| (T - P_{R_{\text{rand}}}^n) r_i \|_R) \]

\[ \| (T - P_{R_{\text{rand}}}^n) \|_{\text{max}} \]
Randomized residual-based error estimators for parametrized equations

(joint work with A. T. Patera and O. Zahm)

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
- Giraldi, Nouy 2017
- Balabanov, Nouy 2018
Goal: Develop a posteriori error estimator for projection-based model order reduction that does not contain constants whose estimation is expensive (inf-sup constant)

Setting: We query a finite number of parameters in the online stage for which we want to estimate the approximation error.

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 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 ∈ ℝᴺ**: random vector such that

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

where Σ 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 \( \|v\|_Σ^2 \)

- 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 \( \|v\|_Σ^2 \)

\[ \frac{1}{K} \sum_{i=1}^{K} (Z_i^T v)^2. \]
**Proposition (Concentration inequality for set of vectors)**

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 $e(\mu_j) = u^N(\mu_j) - u^N(\mu_j)$,

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

that

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

- chi-squared distribution
- concentration around 1 (that means error estimator has perfect effectivity 1)
Proposition (Concentration inequality for set of vectors)

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 $e(\mu_j) = u^N(\mu_j) - u^N(\mu_j)$,

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

that

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

<table>
<thead>
<tr>
<th></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>$#S = 1$</td>
<td>24</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>$#S = 100$</td>
<td>48</td>
<td>16</td>
<td>11</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>$#S = 1000$</td>
<td>60</td>
<td>20</td>
<td>13</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>$#S = 10^6$</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 for set of vectors)

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 $e(\mu_j) = u^N(\mu_j) - u^N(\mu_j)$,

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

that

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

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

- Define solutions of dual problems with random right-hand sides \( Z_i \):
  \[ y_i^N(\mu) := A(\mu)^{-T} Z_i \]

- Approximation of the dual solutions via model order reduction:
  \[ y_i^N(\mu) \approx \tilde{y}_i(\mu) \in \tilde{Y} \subset X^N, \]
  where \( \tilde{Y} \) dual reduced space

- Define fast-to-evaluate randomized error estimator
  \[ \tilde{\Delta}(\mu) := \left( \frac{1}{K} \sum_{i=1}^{K} (\tilde{y}_i(\mu)^T r(\mu))^2 \right)^{1/2} \]
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 $\mu_j$, $j = 1, \ldots, S$ in a finite parameter set $S = \{\mu_1, \ldots, \mu_S\}$ and

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

that

$$\tilde{\Delta}(\mu_j) := \left( \frac{1}{K} \sum_{i=1}^{K} (\tilde{y}_i(\mu_j) Tr(\mu_j))^2 \right)^{1/2}$$

satisfies

$$\mathbb{P}\left\{ (\alpha w)^{-1} \tilde{\Delta}(\mu_j) \leq \| e(\mu_j) \| \Sigma \leq (\alpha w) \tilde{\Delta}(\mu_j), \quad \mu_j \in S \right\} \geq 1 - \delta,$$

where

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

- $\Omega = (0, 1) \times (0, 1)$
- $X = \{ v \in H^1(\Omega) : v(0, x_2) = 0, x_2 \in (0, 1) \}$
- $A(\mu) := -\partial_{x_1} x_1 - \mu_1 \partial x_2 x_2 - \mu_2$
- $P = [0.2, 1.2] \times [10, 50]$
- Neumann b.c. on top: $g_N = \cos(\pi x)$

- high dimensional discretization: linear FE, $h = 0.01$ in each direction
**Histograms of effectivity index** $\tilde{\Delta}(\mu)/\|u(\mu) - u^N(\mu)\|_{H^1(\Omega)}$

5 realizations

- $K = 5$ ($w = 26.1$)
  - $\dim(\tilde{\mathcal{Y}}) = 14 \pm 5$

- $K = 10$ ($w = 6.5$)
  - $\dim(\tilde{\mathcal{Y}}) = 21 \pm 7$

- $K = 20$ ($w = 3.2$)
  - $\dim(\tilde{\mathcal{Y}}) = 28 \pm 8$

**Figure**: $\# \mathcal{S} = 10^4$, $\dim(\mathcal{X}^N) = 20$, vertical dashed lines: $1/w$ and $w$, grey area: $1/(tol \ w)$ and $tol \ w$, where $\alpha \approx tol$, $tol = 2$, solid lines: chi-squared distribution
Randomized a posteriori error estimation

Numerical Experiments

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

100 realizations

Figure: $\#S = 10^4$, $\dim(X^N) = 20$, vertical dashed lines: $1/w$ and $w$, grey area: $1/(tol \ w)$ and $tol \ w$, where $\alpha \approx tol$, $tol = 2$, solid lines: chi-squared distribution.
Summary

- Reduced (local approximation) spaces generated by methods from Randomized Linear Algebra
  - Probabilistic a priori error bound/Numerical experiments: convergence rate is only slightly worse compared to the optimal rate (factor $\sqrt{n}$).
  - Probabilistic a posteriori error bound allows to build the reduced space adaptively
  - required number of local solutions of PDE scale (roughly) with size of the reduced space; Numerical experiments: faster than Lanczos
- Proposed randomized a posteriori error estimator for projection-based model order reduction methods that...
  - ... is based on concentration inequalities, error-residual relationship, and random dual problem
  - ... does only contain computable constants
  - ... is reliable and efficient at high (given) probability
  - ... has a favorable computational complexity as $\dim(\tilde{Y})$ can be chosen relatively small

Thank you very much for your attention!

K Smetana (k.smetana@utwente.nl)
Summary

- Reduced (local approximation) spaces generated by methods from Randomized Linear Algebra
  - Probabilistic a priori error bound/Numerical experiments: convergence rate is only slightly worse compared to the optimal rate (factor $\sqrt{n}$).
  - Probabilistic a posteriori error bound allows to build the reduced space adaptively
  - Required number of local solutions of PDE scale (roughly) with size of the reduced space; Numerical experiments: faster than Lanczos
- Proposed randomized a posteriori error estimator for projection-based model order reduction methods that...
  - ... is based on concentration inequalities, error-residual relationship, and random dual problem
  - ... does only contain computable constants
  - ... is reliable and efficient at high (given) probability
  - ... has a favorable computational complexity as $\dim(\tilde{Y})$ can be chosen relatively small

Thank you very much for your attention!