Model reduction methods for uncertainty quantification

Anthony Nouy

Ecole Centrale Nantes / GeM
Nantes, France

Outline

Uncertainty quantification

Uncertainty quantification has become an essential path in science and engineering

- Assess confidence in numerical predictions
- Comprehension and selection of models (exploration of model predictions over a range of uncertainty, sensitivity analysis)
- Robust optimization and design
- Validation of models from statistical data
- Data assimilation

Uncertainty quantification: stochastic models

Stochastic model

$$u : \theta \in \Theta \mapsto u(\theta) \in \mathcal{V} \quad \text{such that} \quad \mathcal{F}(u(\theta); a(\theta)) = 0$$

where uncertainties are represented by random parameters $a(\theta)$ defined on a probability space $(\Theta, \mathcal{B}, P)$.

- Forward problem:
  For a given probabilistic model for a, compute u or a quantity of interest $Q(u)$ (statistics, probability of events, sensitivity indices...)
- Inverse problem:
  For given observations, estimate a probabilistic model for a.

Uncertainty quantification using functional approaches

- Uncertainties modeled by “simple” random variables

$$\xi : \Theta \rightarrow \mathbb{R}^{N_\xi}, \quad \xi \sim \mu_\xi$$

- Functional representation of any $\sigma(\xi)$-measurable random variable

$$f(\theta) \equiv f(\xi(\theta))$$

- Functional expansions

$$f(\xi) \approx \sum_{j=1}^{N_\eta} \eta_j \psi_j(\xi) := f(\xi; \eta)$$

Stochastic (parametric) analyses

$$\text{Random parameters } a(\xi) = a(\xi;\eta), \quad \xi \sim \mu_\xi$$

- Forward problem : given μ_ξ and η, compute $\mathcal{O}(u)$ with

$$\mathcal{F}(u(\xi); a(\xi;\eta)) = 0$$

Stochastic (parametric) analyses

$$\text{Random parameters } a(\xi) = a(\xi; \eta), \quad \xi \sim \mu_\xi$$

- Forward problem : given μ_ξ and η, compute $\mathcal{O}(u)$ with

$$\mathcal{F}(u(\xi); a(\xi; \eta)) = 0$$

- Inverse problem: given $\mathcal{O}(u)$, determine $\eta \in \Xi_\eta \subset \mathbb{R}^{N_\eta}$ (or μ_η) with

$$\mathcal{F}(u(\xi, \eta); a(\xi; \eta)) = 0$$

Example: Diffusion in heterogeneous media

Stochastic PDE with a random diffusion matrix a:

$$-\nabla \cdot (a\nabla u) = f \quad + \quad \text{boundary conditions}$$

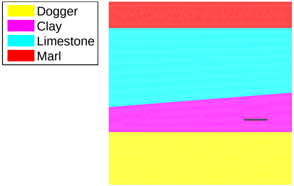
Example: Diffusion in heterogeneous media

Stochastic PDE with a random diffusion matrix a:

$$-\nabla \cdot (a \nabla u) = f \quad + \quad \text{boundary conditions}$$

- Groundwater flow.
  Geological layers with uncertain properties:

$$a(x, \theta) = \sum_{i=1}^{4} \xi_i(\theta) I_{\Omega_i}(x)$$



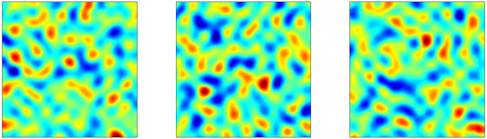
ξ_i's probability laws

Layer	Law
Dogger	$LU(5, 125)$
Clay	$LU(3.10^{-7}, 3.10^{-5})$
Limestone	$LU(1.2, 30)$
Marl	$LU(10^{-5}, 10^{-4})$

- Random media with spatially correlated random fields

$$a(x,\theta) = \underline{a}(x) + \sum_{i=1}^{m} \sqrt{\sigma_i}\varphi_i(x)\lambda_i(\xi(\theta)), \quad \xi = (\xi_1, \ldots, \xi_{N_\xi})$$

$$\lambda_i(\xi) = \sum_{\alpha} \eta_{i,\alpha}\psi_\alpha(\xi)$$



Uncertainty quantification using functional approaches

Ideal approach

Compute an accurate and explicit approximation of $u(\xi)$ (or $u(\xi, \eta)$) (metamodel, surrogate...) that allows fast evaluations of output quantities of interest, observables, or objective function.

Issue

Approximation of a high dimensional function

$$u(x), \; x \in \Xi \subset \mathbb{R}^d,$$

with $(\Xi, \mathcal{B}(\Xi), \mu)$ a measured space.

Outline

1. Uncertainty quantification
2. Model reduction methods for high dimensional problems
3. Sparse approximation
4. Tensor-structured problems and low-rank tensor approximation
5. Numerical methods for equations in tensor format
6. Low-rank approximation and subspace-based model reduction
7. Subspace-based model reduction for higher-order tensors

High dimensional problems

- Many problems in computational science requires the approximation of high dimensional functions:
$$u(x_1, \ldots, x_d)$$
- Classical discretization methods introduce high-dimensional parametrizations
$$u(x_1, \ldots, x_d) \approx \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} a_{i_1 \ldots i_d} \varphi^1_{i_1}(x_1) \ldots \varphi^1_{i_d}(x_d)$$
$$a \in \mathbb{R}^{n_1 \times \ldots \times n_d}$$
- Model order reduction methods aim at replacing a complex model with a simplified one, living in a lower dimensional space (or manifold).

Model reduction methods

- Order reduction methods exploit specific structures (application dependent)
  - Smoothness
  - Low effective dimensionality, e.g.

$$u(x_1, \dots, x_d) \approx g(x_1, x_2)$$

  - Low-order interactions, e.g.

$$u(x_1, \dots, x_d) \approx u_0 + \sum_i u_i(x_i) + \sum_{i \neq j} u_{i,j}(x_i, x_j)$$

  - Sparsity (relatively to a basis or frame)
  - Low-rank structure

- Structures possibly discovered with suitable parametrizations

$$u(x_1, \dots, x_d) \approx g(y_1, \dots, y_m), \quad (y_1, \dots, y_m) = h(x_1, \dots, x_d),$$

with g smooth, sparse, low-rank, ...

Model reduction methods

- Order reduction methods exploit specific structures (application dependent)
  - Smoothness
  - Low effective dimensionality, e.g.

$$u(x_1, \dots, x_d) \approx g(x_1, x_2)$$

  - Low-order interactions, e.g.

$$u(x_1, \dots, x_d) \approx u_0 + \sum_i u_i(x_i) + \sum_{i \neq j} u_{i,j}(x_i, x_j)$$

  - Sparsity (relatively to a basis or frame)
  - Low-rank structure

- Structures possibly discovered with suitable parametrizations

$$u(x_1, \dots, x_d) \approx g(y_1, \dots, y_m), \quad (y_1, \dots, y_m) = h(x_1, \dots, x_d),$$

with g smooth, sparse, low-rank, ...

Model reduction methods

- Depend on the objectives
  - L^∞ optimality (global optimization, inverse problems)
  - L^p optimality (energy, moments)
  - ...
- Depends on the available information on the function
  - Pointwise evaluations
$$u(x_1^k, \ldots, x_d^k)$$
  - Equations (ODE, PDE, DAE)
  - Partial pointwise evaluations (e.g. for parametric/stochastic problems): equations for
$$u(\cdot, x_2^k, \ldots, x_d^k)$$

Model reduction methods

- Depend on the objectives
  - L^∞ optimality (global optimization, inverse problems)
  - L^p optimality (energy, moments)
  - ...
- Depends on the available information on the function
  - Pointwise evaluations
  $$u(x_1^k, \dots, x_d^k)$$
  - Equations (ODE, PDE, DAE)
  - Partial pointwise evaluations (e.g. for parametric/stochastic problems): equations for
  $$u(\cdot, x_2^k, \dots, x_d^k)$$

Outline

Sparse approximation

- We want to approximate a function $u(x)$ in a certain vector space V equipped with a norm $\|\cdot\|$. Consider a set of functions $\{\psi_\alpha; \alpha \in \Lambda\}$ (called a dictionary) whose span is dense in V, so that

$$u(x) = \sum_{\alpha \in \Lambda} u_\alpha \psi_\alpha(x)$$

Typically, for multivariate functions

$$\psi_\alpha(x) = \psi^1_{\alpha_1}(x_1) \ldots \psi^d_{\alpha_d}(x_d), \quad \Lambda = \left\{\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d; \sum_{i=1}^d \alpha_i < \infty\right\}$$

- Sparse approximation methods rely on the fact that a good approximation (or even an exact decomposition) of the solution can be obtained by only considering a small subset of functions:

$$u(x) \approx u_n(x) = \sum_{\alpha \in \Lambda_n} u_\alpha \psi_\alpha(x), \quad \#\Lambda_n = n$$

Best n-term approximation

The best n-term approximation is the solution of

$$\inf_{u_n \in \mathcal{M}_n} \|u - u_n\| := \sigma_n(u)$$

where

$$\mathcal{M}_n = \left\{ v = \sum_{\alpha \in \Lambda_n} v_\alpha \psi_\alpha ; v_\alpha \in \mathbb{R}; \Lambda_n \subset \Lambda, \ \#\Lambda_n = n \right\}$$

- Selection of basis (or frame) suitable for a class of functions $\mathcal{K}$?

$$\sup_{u \in \mathcal{K}} \sigma_n(u) \leq c(n, \mathcal{K}) \underset{n \to \infty}{\longrightarrow} 0 \quad \text{with fast decay}$$

- Best n-term approximation is a combinatorial problem.
- Computation of the (quasi)-best n-term approximation through adaptive or non-adaptive algorithms ?

Sparse approximation: alternative strategies

- A priori definition of an index set $\widetilde{\Lambda}$, e.g. based on the properties of the class of functions we want to approximate (regularity, anisotropy...).
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \alpha_i \leq p\}$
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \gamma_i \alpha_i \leq p\}$ (anisotropic)
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d g(\alpha_i) \leq g(p)\}$ (smolyak)
  - ...

  A priori error estimation can be useful.
- Adaptive construction based on a posteriori error estimation, e.g. by greedy algorithms:

$$\Lambda_{n+1} = \Lambda_n \cup \{\alpha_{n+1}\}, \quad \text{with } \alpha_{n+1} \text{ in a candidate set.}$$

- Non-adaptive approximation by solving (approximately) the best n-term approximation problem.

$$\mathcal{M}_n = \{\sum_{\alpha\in\Lambda} v_\alpha \psi_\alpha ; \|\mathbf{v}\|_0 \leq n\} \xrightarrow{\text{convexification}} \mathcal{M}^\gamma = \{\sum_{\alpha\in\Lambda} v_\alpha \psi_\alpha ; \|\mathbf{v}\|_1 \leq \gamma\}$$

Sparse approximation: alternative strategies

- A priori definition of an index set $\widetilde{\Lambda}$, e.g. based on the properties of the class of functions we want to approximate (regularity, anisotropy...).
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \alpha_i \leq p\}$
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \gamma_i \alpha_i \leq p\}$ (anisotropic)
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d g(\alpha_i) \leq g(p)\}$ (smolyak)
  - ...

  A priori error estimation can be useful.
- Adaptive construction based on a posteriori error estimation, e.g. by greedy algorithms:

$$\Lambda_{n+1} = \Lambda_n \cup \{\alpha_{n+1}\}, \quad \text{with } \alpha_{n+1} \text{ in a candidate set.}$$

- Non-adaptive approximation by solving (approximately) the best n-term approximation problem.

$$\mathcal{M}_n = \{\sum_{\alpha \in \Lambda} v_\alpha \psi_\alpha; \|\mathbf{v}\|_0 \leq n\} \xrightarrow{\text{convexification}} \mathcal{M}^\gamma = \{\sum_{\alpha \in \Lambda} v_\alpha \psi_\alpha; \|\mathbf{v}\|_1 \leq \gamma\}$$

Sparse approximation: alternative strategies

- A priori definition of an index set $\widetilde{\Lambda}$, e.g. based on the properties of the class of functions we want to approximate (regularity, anisotropy...).
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d ; \sum_{i=1}^d \alpha_i \leq p\}$
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d ; \sum_{i=1}^d \gamma_i \alpha_i \leq p\}$ (anisotropic)
  - $\widetilde{\Lambda} = \{\alpha \in \mathbb{N}^d ; \sum_{i=1}^d g(\alpha_i) \leq g(p)\}$ (smolyak)
  - ...

  A priori error estimation can be useful.

- Adaptive construction based on a posteriori error estimation, e.g. by greedy algorithms:

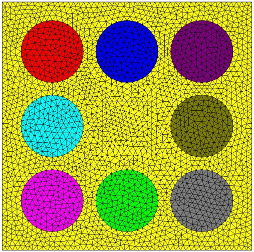
$$\Lambda_{n+1} = \Lambda_n \cup \{\alpha_{n+1}\}, \quad \text{with } \alpha_{n+1} \text{ in a candidate set.}$$

- Non-adaptive approximation by solving (approximately) the best n-term approximation problem.

$$\mathcal{M}_n = \{\sum_{\alpha\in\Lambda} v_\alpha \psi_\alpha ; \|\mathbf{v}\|_0 \leq n\} \xrightarrow{\text{convexification}} \mathcal{M}^\gamma = \{\sum_{\alpha\in\Lambda} v_\alpha \psi_\alpha ; \|\mathbf{v}\|_1 \leq \gamma\}$$

Illustration: diffusion problem with multiple inclusions

$$\begin{cases} -\nabla \cdot (\kappa \nabla u) = I_D(x) \quad on \quad \Omega = (0,1)\times(0,1) \\ u = 0 \quad on \quad \partial\Omega \end{cases}$$



with

$$\kappa(x,\xi) = \begin{cases} 1 & \text{if } x \in \Omega_0 \\ 1 + 0.1\xi_i & \text{if } x \in \Omega_i,\ i = 1...8 \end{cases}$$

with $\xi_i \sim U(-1,1)$. $\Xi = (-1,1)^8$.

- Quantity of Interest:

$$I(\xi) = \int_D u(x,\xi)dx, \quad D = (0.4,0.6)\times(0.4,0.6)$$

- Polynomial approximation:

$$I(\xi) \approx \sum_{\alpha\in\Lambda} v_\alpha \psi_\alpha(\xi), \quad span\{\psi_\alpha : \alpha \in \Lambda\} = \mathbb{P}_4(\Xi), \quad dim(\mathbb{P}_4(\Xi)) = 1286$$

Coefficients $\{v_\alpha : \alpha \in \Lambda\}$ computed with least-squares with K samples.

$$\min_{\{I_\alpha\}} \sum_{k=1}^{K} (I(\xi^k) - \sum_{\alpha} v_\alpha \psi_\alpha(\xi^k))^2$$

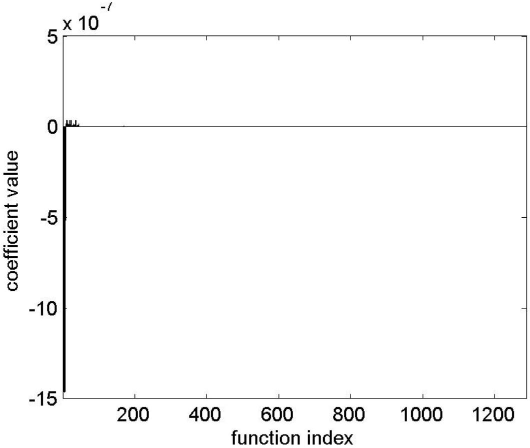


Figure : Values of the 1286 coefficients when using a large number of samples $K = 2000$

Coefficients $\{v_\alpha : \alpha \in \Lambda\}$ computed with least-squares with K samples.

$$\min_{\{I_\alpha\}} \sum_{k=1}^{K} (I(\xi^k) - \sum_\alpha v_\alpha \psi_\alpha(\xi^k))^2$$

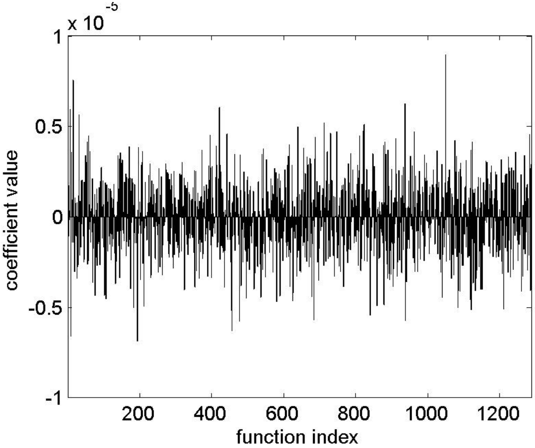


Figure : Values of the 1286 coefficients when using a small number of samples $Q = 50$

Least-squares with sparcity-inducing regularization (ℓ_1 regularization):

$$\min_{\{I_\alpha\}} \sum_{k=1}^{K} (I(\xi^k) - \sum_\alpha v_\alpha \psi_\alpha(\xi^k))^2 + \sum_\alpha |v_\alpha|$$

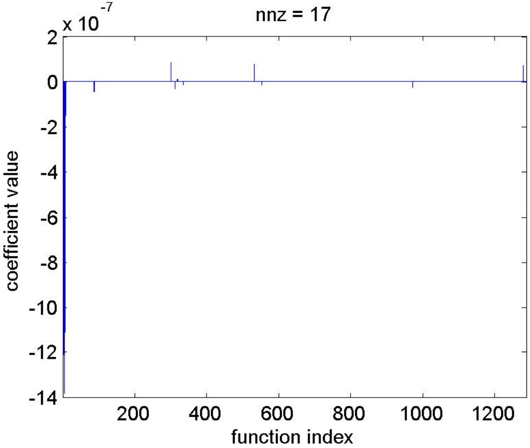


Figure : Values of the 1286 coefficients when using a small number of samples $K = 50$

Outline

1. Uncertainty quantification
2. Model reduction methods for high dimensional problems
3. Sparse approximation
4. Tensor-structured problems and low-rank tensor approximation
5. Numerical methods for equations in tensor format
6. Low-rank approximation and subspace-based model reduction
7. Subspace-based model reduction for higher-order tensors

Low rank tensor approximation :

What is a tensor ?

What is the rank of a tensor ?

Tensor spaces

An algebraic tensor space $V = V_1 \otimes_a \ldots \otimes_a V_d$ is the set of elements of the form

$$u = \sum_{i=1}^{m} v_i^1 \otimes \ldots \otimes v_i^d$$

or for multivariate functions

$$u(x_1, \ldots, x_d) = \sum_{i=1}^{m} v_i^1(x_1) \ldots v_i^d(x_d).$$

A tensor Banach space $V_{\|\cdot\|}$ is obtained by the completion of the algebraic tensor space V with respect to a norm $\|\cdot\|$:

$$V_{\|\cdot\|} = \overline{V_1 \otimes_a \ldots \otimes_a V_d}^{\|\cdot\|}.$$

W. Hackbusch.
Tensor Spaces and Numerical Tensor Calculus,
Springer, 2012.

Examples of (Banach) tensor spaces

- Finite dimensional tensor spaces: For finite dimensional spaces,

$$V_{\|\cdot\|} = V = V_1 \otimes_a \ldots \otimes_a V_d$$

Denoting $\{\phi_i^k\}_{i=1}^{n_k}$ a basis of the n_k-dimensional space V_k, $u \in V_{\|\cdot\|}$ can be written

$$u = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} a_{i_1 \ldots i_d} \phi_{i_1}^1 \otimes \ldots \otimes \phi_{i_d}^d,$$

and identified with

$$a \in \mathbb{R}^{n_1 \times \ldots \times n_d} = \mathbb{R}^{n_1} \otimes \ldots \otimes \mathbb{R}^{n_d}$$

Examples of (Banach) tensor spaces

- Bochner space $L^p_\mu(I;\mathcal{V})$, the set of Bochner measurable functions u defined on a measure space (I,μ) with values in a Banach space $(\mathcal{V},\|\cdot\|_\mathcal{V})$, with bounded norm

$$\|u\|_p = \left(\int_I \|u(y)\|^p_\mathcal{V}\mu(dy)\right)^{1/p} \qquad (1 \le p < \infty),$$

$$\text{or} \quad \|u\|_\infty = \operatorname*{ess\,sup}_{y\in I} \|u(y)\|_\mathcal{V} \qquad (p = \infty)$$

- An element $u \in L^p_\mu(I) \otimes_a \mathcal{V}$ is of the form

$$u(y) = \sum_{i=1}^m w_i \lambda_i(y), \quad y \in I.$$

- Case $1 \le p < \infty$.

$$\overline{L^p_\mu(I) \otimes_a \mathcal{V}}^{\|\cdot\|_p} = L^p_\mu(I;\mathcal{V})$$

- Case $p = \infty$.

$$\overline{L^\infty_\mu(I) \otimes_a \mathcal{V}}^{\|\cdot\|_\infty} \subset L^\infty_\mu(I;\mathcal{V})$$

Examples of (Banach) tensor spaces

- Bochner space $L^p_\mu(I;\mathcal{V})$, the set of Bochner measurable functions u defined on a measure space (I,μ) with values in a Banach space $(\mathcal{V},\|\cdot\|_\mathcal{V})$, with bounded norm

$$\|u\|_p = \left(\int_I \|u(y)\|^p_\mathcal{V}\mu(dy)\right)^{1/p} \qquad (1\leq p<\infty),$$

$$\text{or} \quad \|u\|_\infty = \operatorname*{ess\,sup}_{y\in I}\|u(y)\|_\mathcal{V} \qquad (p=\infty)$$

  - An element $u\in L^p_\mu(I)\otimes_a\mathcal{V}$ is of the form

$$u(y)=\sum_{i=1}^m w_i\lambda_i(y), \quad y\in I.$$

  - Case $1\leq p<\infty$.

$$\overline{L^p_\mu(I)\otimes_a\mathcal{V}}^{\|\cdot\|_p} = L^p_\mu(I;\mathcal{V})$$

  - Case $p=\infty$.

$$\overline{L^\infty_\mu(I)\otimes_a\mathcal{V}}^{\|\cdot\|_\infty} \subset L^\infty_\mu(I;\mathcal{V})$$

Examples of (Banach) tensor spaces

- Bochner space $L^p_\mu(I;\mathcal{V})$, the set of Bochner measurable functions u defined on a measure space (I,μ) with values in a Banach space $(\mathcal{V},\|\cdot\|_\mathcal{V})$, with bounded norm

$$\|u\|_p = \left(\int_I \|u(y)\|_\mathcal{V}^p \mu(dy)\right)^{1/p} \qquad (1 \le p < \infty),$$

$$\text{or} \quad \|u\|_\infty = \operatorname*{ess\,sup}_{y\in I} \|u(y)\|_\mathcal{V} \qquad (p=\infty)$$

  - An element $u \in L^p_\mu(I) \otimes_a \mathcal{V}$ is of the form

$$u(y) = \sum_{i=1}^m w_i \lambda_i(y), \quad y \in I.$$

  - Case $1 \le p < \infty$.

$$\overline{L^p_\mu(I) \otimes_a \mathcal{V}}^{\|\cdot\|_p} = L^p_\mu(I;\mathcal{V})$$

  - Case $p = \infty$.

$$\overline{L^\infty_\mu(I) \otimes_a \mathcal{V}}^{\|\cdot\|_\infty} \subset L^\infty_\mu(I;\mathcal{V})$$

Examples of (Banach) tensor spaces

- Bochner space $L^p_\mu(I;\mathcal{V})$, the set of Bochner measurable functions u defined on a measure space (I,μ) with values in a Banach space $(\mathcal{V},\|\cdot\|_\mathcal{V})$, with bounded norm

$$\|u\|_p = \left(\int_I \|u(y)\|^p_\mathcal{V}\mu(dy)\right)^{1/p} \qquad (1\leq p<\infty),$$

$$\text{or} \quad \|u\|_\infty = \operatorname*{ess\,sup}_{y\in I}\|u(y)\|_\mathcal{V} \qquad (p=\infty)$$

  - An element $u\in L^p_\mu(I)\otimes_a\mathcal{V}$ is of the form

$$u(y)=\sum_{i=1}^m w_i\lambda_i(y),\quad y\in I.$$

  - Case $1\leq p<\infty$.

$$\overline{L^p_\mu(I)\otimes_a\mathcal{V}}^{\|\cdot\|_p} = L^p_\mu(I;\mathcal{V})$$

  - Case $p=\infty$.

$$\overline{L^\infty_\mu(I)\otimes_a\mathcal{V}}^{\|\cdot\|_\infty} \subset L^\infty_\mu(I;\mathcal{V})$$

Examples of (Banach) tensor spaces

- Lebesgue spaces $L^p_\mu(I)$ with product measures $\mu = \mu_1 \otimes \ldots \otimes \mu_d$ on $I = I_1 \times \ldots \times I_d$:

$$L^p_\mu(I_1 \times \ldots \times I_d) = \overline{L^p_{\mu_1}(I_1) \otimes_a \ldots \otimes_a L^p_{\mu_1}(I_d)}^{\|\cdot\|_p} \quad (1 \leq p < \infty)$$

Examples of (Banach) tensor spaces

- Lebesgue spaces $L^p_\mu(I)$ with product measures $\mu = \mu_1 \otimes \ldots \otimes \mu_d$ on $I = I_1 \times \ldots \times I_d$:

$$L^p_\mu(I_1 \times \ldots \times I_d) = \overline{L^p_{\mu_1}(I_1) \otimes_a \ldots \otimes_a L^p_{\mu_1}(I_d)}^{\|\cdot\|_p} \quad (1 \leq p < \infty)$$

- Sobolev spaces $W^{s,p}(I)$ on $I = I_1 \times \ldots \times I_d$, the set of measurable functions $u : I \to \mathbb{R}$ with bounded norm

$$\|u\|_{s,p} = \sum_{|\alpha| \leq s} \|\partial^\alpha u\|_p, \quad \partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d}.$$

$$W^{s,p}(I) = \overline{W^{s,p}(I_1) \otimes_a \ldots \otimes_a W^{s,p}(I_d)}^{\|\cdot\|_{s,p}} \quad (1 \leq p < \infty)$$

$W^{s,p}(I)$ is an intersection tensor space:

$$W^{s,p}(I) = \bigcap_{\alpha \in \Lambda_s} \overline{W^{\alpha_1,p} \otimes_a \ldots \otimes_a W^{\alpha_d,p}}^{\|\cdot\|_\alpha}$$

$$\Lambda_s = \{(0, \ldots, 0), (s, 0, \ldots, 0), (0, \ldots, 0, s)\}$$

Tensor-structured problems in Uncertainty Quantification

- Stochastic/Parametric equations (PDEs, ODEs...):

$$A_\xi(u(\xi)) = b_\xi, \quad u(\xi) \in \mathcal{V}$$

$$\xi \sim \mu, \quad supp(\mu) = \Xi$$

with parameter dependent operator A_ξ and right-hand side b_ξ.

$$u \in L^p_\mu(\Xi; \mathcal{V}) = \mathcal{V} \otimes L^p_\mu(\Xi)$$

- Functions of independent random variables:

$$u(\xi_1, \xi_2, \ldots, \xi_d)$$

$$\xi_k \sim \mu_k, \quad supp(\mu_k) = \Xi_k$$

$$u \in \mathcal{V} \otimes L^p_\mu(\Xi) = \mathcal{V} \otimes L^p_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^p_{\mu_d}(\Xi_d)$$

Tensor-structured problems in Uncertainty Quantification

- Stochastic/Parametric equations (PDEs, ODEs...):

$$A_\xi(u(\xi)) = b_\xi, \quad u(\xi) \in \mathcal{V}$$

$$\xi \sim \mu, \quad supp(\mu) = \Xi$$

with parameter dependent operator A_ξ and right-hand side b_ξ.

$$u \in L^p_\mu(\Xi; \mathcal{V}) = \mathcal{V} \otimes L^p_\mu(\Xi)$$

- Functions of independent random variables:

$$u(\xi_1, \xi_2, \dots, \xi_d)$$

$$\xi_k \sim \mu_k, \quad supp(\mu_k) = \Xi_k$$

$$u \in \mathcal{V} \otimes L^p_\mu(\Xi) = \mathcal{V} \otimes L^p_{\mu_1}(\Xi_1) \otimes \dots \otimes L^p_{\mu_d}(\Xi_d)$$

- Parametrized functions of random variables (robust optimization and control, statistical inverse problems):

$$u(\xi, \eta)$$

$$\xi \sim \mu, \quad supp(\mu) = \Xi, \quad \eta \in A$$

$$u \in \mathcal{V} \otimes L^p_\mu(\Xi) \otimes L^q_\nu(A)$$

- Stochastic calculus:

$$\begin{cases} dX_t = a(X_t, t)dt + \sigma(X_t, t)dW_t \\ X_0 = x_0 \end{cases} \qquad X_t = (X_t^1 \ldots X_t^n)$$

The probability density function $u(\cdot, t)$ of X_t verifies a n-dimensional PDE (Kolmogorov forward equation)

$$u(\cdot, t) \in H^1_{\mu_1}(\mathbb{R}) \otimes \ldots \otimes H^1_{\mu_n}(\mathbb{R})$$

- Parametrized functions of random variables (robust optimization and control, statistical inverse problems):

$$u(\xi, \eta)$$

$$\xi \sim \mu, \quad supp(\mu) = \Xi, \quad \eta \in A$$

$$u \in \mathcal{V} \otimes L^p_\mu(\Xi) \otimes L^q_\nu(A)$$

- Stochastic calculus:

$$\begin{cases} dX_t = a(X_t, t)dt + \sigma(X_t, t)dW_t \\ X_0 = x_0 \end{cases} \qquad X_t = (X_t^1 \dots X_t^n)$$

The probability density function $u(\cdot, t)$ of X_t verifies a n-dimensional PDE (Kolmogorov forward equation)

$$u(\cdot, t) \in H^1_{\mu_1}(\mathbb{R}) \otimes \dots \otimes H^1_{\mu_n}(\mathbb{R})$$

Other tensor structured problems in computational science

- Dynamical systems:

$$\partial_t u(t) = F(u(t); t)$$

$$t \in I, \quad u(t) \in \mathcal{V}$$

$$u \in L^p(I) \otimes \mathcal{V}$$

- Multidimensional PDEs

$$\left(\sum_{i,j=1}^{d} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} + c \right) u(x_1, \ldots, x_d) = f$$

$$(x_1, \ldots, x_d) \in \Omega_1 \times \ldots \times \Omega_d$$

$$u \in H^1(\Omega_1) \otimes \ldots \otimes H^1(\Omega_d)$$

- Quantum physics and chemistry (Schrödinger equation, Master equation, ...)
- Problem in a vector space can be made a problem in a tensor space, e.g. through quantization

$$u \in \mathbb{R}^{2^d} \simeq \mathbb{R}^2 \otimes \ldots \otimes \mathbb{R}^2$$

Other tensor structured problems in computational science

- Dynamical systems:

$$\partial_t u(t) = F(u(t); t)$$

$$t \in I, \quad u(t) \in \mathcal{V}$$

$$u \in L^p(I) \otimes \mathcal{V}$$

- Multidimensional PDEs

$$\left(\sum_{i,j=1}^{d} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} + c \right) u(x_1, \ldots, x_d) = f$$

$$(x_1, \ldots, x_d) \in \Omega_1 \times \ldots \times \Omega_d$$

$$u \in H^1(\Omega_1) \otimes \ldots \otimes H^1(\Omega_d)$$

- Quantum physics and chemistry (Schrödinger equation, Master equation, ...)
- Problem in a vector space can be made a problem in a tensor space, e.g. through quantization

$$u \in \mathbb{R}^{2^d} \simeq \mathbb{R}^2 \otimes \ldots \otimes \mathbb{R}^2$$

Other tensor structured problems in computational science

- Dynamical systems:

$$\partial_t u(t) = F(u(t); t)$$
$$t \in I, \quad u(t) \in \mathcal{V}$$
$$u \in L^p(I) \otimes \mathcal{V}$$

- Multidimensional PDEs

$$\left(\sum_{i,j=1}^{d} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} + c \right) u(x_1, \ldots, x_d) = f$$
$$(x_1, \ldots, x_d) \in \Omega_1 \times \ldots \times \Omega_d$$
$$u \in H^1(\Omega_1) \otimes \ldots \otimes H^1(\Omega_d)$$

- Quantum physics and chemistry (Schrödinger equation, Master equation, ...)
- Problem in a vector space can be made a problem in a tensor space, e.g. through quantization

$$u \in \mathbb{R}^{2^d} \simeq \mathbb{R}^2 \otimes \ldots \otimes \mathbb{R}^2$$

Other tensor structured problems in computational science

- Dynamical systems:

$$\partial_t u(t) = F(u(t); t)$$

$$t \in I, \quad u(t) \in \mathcal{V}$$

$$u \in L^p(I) \otimes \mathcal{V}$$

- Multidimensional PDEs

$$\left(\sum_{i,j=1}^{d} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} + c \right) u(x_1, \ldots, x_d) = f$$

$$(x_1, \ldots, x_d) \in \Omega_1 \times \ldots \times \Omega_d$$

$$u \in H^1(\Omega_1) \otimes \ldots \otimes H^1(\Omega_d)$$

- Quantum physics and chemistry (Schrödinger equation, Master equation, ...)
- Problem in a vector space can be made a problem in a tensor space, e.g. through quantization

$$u \in \mathbb{R}^{2^d} \simeq \mathbb{R}^2 \otimes \ldots \otimes \mathbb{R}^2$$

Low-rank approximation

- Approximation in a subset of tensors with bounded rank:

$$\mathcal{M}_{\leq r} = \{v \in V = V_1 \otimes \ldots \otimes V_d ; rank(v) \leq r\}$$

- For order-two tensors, a single notion of rank.

$$v \in V_1 \otimes V_2$$

$$rank(v) \leq r \iff v = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \quad \text{or} \quad v(x) = \sum_{i=1}^{r} v_i^1(x_1) v_i^2(x_2)$$

- For higher-order tensors, different notions of rank.

Low-rank approximation

- Approximation in a subset of tensors with bounded rank:

$$\mathcal{M}_{\leq r} = \{v \in V = V_1 \otimes \ldots \otimes V_d ; rank(v) \leq r\}$$

- For order-two tensors, a single notion of rank.

$$v \in V_1 \otimes V_2$$

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \quad \text{or} \quad v(x) = \sum_{i=1}^{r} v_i^1(x_1) v_i^2(x_2)$$

- For higher-order tensors, different notions of rank.

Low-rank approximation

- Approximation in a subset of tensors with bounded rank:

$$\mathcal{M}_{\leq r} = \{v \in V = V_1 \otimes \ldots \otimes V_d ; rank(v) \leq r\}$$

- For order-two tensors, a single notion of rank.

$$v \in V_1 \otimes V_2$$

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \quad \text{or} \quad v(x) = \sum_{i=1}^{r} v_i^1(x_1) v_i^2(x_2)$$

- For higher-order tensors, different notions of rank.

- Canonical rank:

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes \ldots \otimes v_i^d \quad \text{or} \quad v(x) = \sum_{i=1}^{r} v_i^1(x_1) \ldots v_i^d(x_d)$$

- α-rank:
  for $\alpha \subset \{1, \ldots, d\}$, $V = V_\alpha \otimes V_{\alpha^c}$, with $V_\alpha = \bigotimes_{\mu \in \alpha} V_\mu$ and define the α-rank:

$$rank_\alpha(v) \leq r_\alpha \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r_\alpha} v_i^\alpha \otimes v_i^{\alpha^c}, \quad v_i^\alpha \in V_\alpha, \quad v_i^{\alpha^c} \in V_{\alpha^c}$$

Example:

$u(x_1, \ldots, x_4) = f(x_1, x_2) g(x_3, x_4)$ is such that $rank_{(1,2)}(u) = rank_{(3,4)}(u) = 1$.

Higher-order low-rank formats

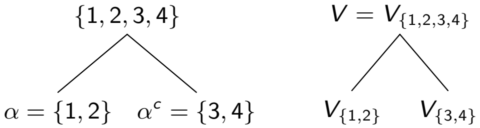
- Canonical rank:

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes \ldots \otimes v_i^d \quad \text{or} \quad v(x) = \sum_{i=1}^{r} v_i^1(x_1) \ldots v_i^d(x_d)$$

- α-rank:
for $\alpha \subset \{1, \ldots, d\}$, $V = V_\alpha \otimes V_{\alpha^c}$, with $V_\alpha = \bigotimes_{\mu \in \alpha} V_\mu$ and define the α-rank:

$$rank_\alpha(v) \leq r_\alpha \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r_\alpha} v_i^\alpha \otimes v_i^{\alpha^c}, \quad v_i^\alpha \in V_\alpha, \quad v_i^{\alpha^c} \in V_{\alpha^c}$$

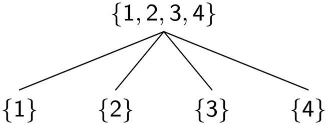
Example:



$u(x_1, \ldots, x_4) = f(x_1, x_2) g(x_3, x_4)$ is such that $rank_{(1,2)}(u) = rank_{(3,4)}(u) = 1$.

- Tucker rank:

$$rank_T(v) = (rank_1(v), \dots, rank_d(v))$$



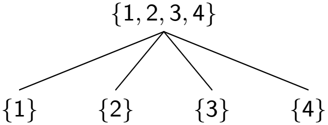
$$rank_T(v) \leq r = (r_1, \dots, r_d) \iff v = \sum_{i_1=1}^{r_1} \dots \sum_{i_d=1}^{r_d} a_{i_1 \dots i_d} v^1_{i_1} \otimes \dots \otimes v^d_{i_d}$$

- Tree-based Tucker rank:

$$rank_T(v) = (rank_\alpha(v); \alpha \in T) \quad \text{with } T \text{ a dimension tree}$$

- Tucker rank:

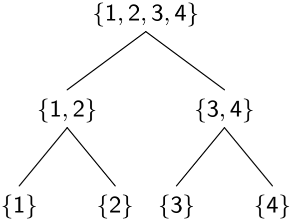
$$rank_T(v) = (rank_1(v), \dots, rank_d(v))$$



$$rank_T(v) \leq r = (r_1, \dots, r_d) \iff v = \sum_{i_1=1}^{r_1} \dots \sum_{i_d=1}^{r_d} a_{i_1 \dots i_d} v_{i_1}^1 \otimes \dots \otimes v_{i_d}^d$$

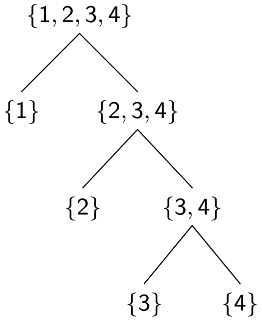
- Tree-based Tucker rank:

$$rank_T(v) = (rank_\alpha(v); \alpha \in T) \quad \text{with } T \text{ a dimension tree}$$



- TT-rank:

$$rank_{TT}(v) = (rank_1(v), \dots, rank_{d-1}(v))$$



$$rank_{TT}(v) \leq r = (r_1, \dots, r_{d-1}) \quad \Longleftrightarrow \quad v = \sum_{i_1=1}^{r_1} \dots \sum_{i_d=1}^{r_{d-1}} v^1_{i_1,i_2} \otimes \dots \otimes v^d_{i_{d-1},1}$$

Low-rank tensor subsets

- Different notions of rank yield different low-rank tensor subsets: Canonical, Tucker, Tree-based Tucker (HT, TT), ...

$$\mathcal{M}_{\leq r} = \{v \in V : rank(v) \leq r\}$$

- $\mathcal{M}_{\leq r}$ has a small dimension $n(r,d)$ (i.e. can be parameterized with a small number n of parameters), typically

$$n(r,d) = O(dr^s)$$

Low-rank tensor subsets

- Different notions of rank yield different low-rank tensor subsets: Canonical, Tucker, Tree-based Tucker (HT, TT), ...

$$\mathcal{M}_{\leq r} = \{v \in V : rank(v) \leq r\}$$

- $\mathcal{M}_{\leq r}$ has a small dimension $n(r,d)$ (i.e. can be parameterized with a small number n of parameters), typically

$$n(r,d) = O(dr^s)$$

Approximation in low-rank tensor subsets

- Good approximation properties: for a large set of functions $C(V)$ of interest,

$$\inf_{v \in \mathcal{M}} \|u - v\| \leq \epsilon(n) \quad \forall u \in C(V)$$

  with rapidly decaying $\epsilon(n)$.

- Good approximation for smooth functions

Example (Sobolev regularity: approximation in canonical format)

$$\inf_{v \in \mathcal{R}_r} \|u - v\|_{L^2} \lesssim r^{-sd/(d-1)} \quad \forall u \in B_s^{mix} \subset L^2(\pi_d)$$

$$\text{with} \quad B_s^{mix} = \left\{ u \in L^2(\pi_d); \|u\|_{H_s^{mix}} \leq 1 \right\} \subset H_s^{mix}(\pi_d)$$

That means that for any $u \in B_s^{mix}$ and for $\epsilon > 0$, it could be possible to find an approximation $v(x_1, \ldots, x_d) = \sum_{i=1}^r \phi_i^1(x_1) \ldots \phi_i^d(x_d)$ such that

$$\|u - v\| \leq \epsilon \quad \text{with} \quad r \gtrsim \epsilon^{-\frac{d-1}{sd}}$$

- But low-rank approximations are expected to exploit additional features.

Approximation in low-rank tensor subsets

- Good approximation properties: for a large set of functions $C(V)$ of interest,

$$\inf_{v\in\mathcal{M}} \|u-v\| \leq \epsilon(n) \quad \forall u \in C(V)$$

with rapidly decaying $\epsilon(n)$.

- Good approximation for smooth functions

Example (Sobolev regularity: approximation in canonical format)

$$\inf_{v\in\mathcal{R}_r} \|u-v\|_{L^2} \lesssim r^{-sd/(d-1)} \quad \forall u \in B_s^{mix} \subset L^2(\pi_d)$$

$$\text{with} \quad B_s^{mix} = \left\{u \in L^2(\pi_d); \|u\|_{H_s^{mix}} \leq 1\right\} \subset H_s^{mix}(\pi_d)$$

That means that for any $u \in B_s^{mix}$ and for $\epsilon > 0$, it could be possible to find an approximation $v(x_1,\ldots,x_d) = \sum_{i=1}^r \phi_i^1(x_1)\ldots\phi_i^d(x_d)$ such that

$$\|u-v\| \leq \epsilon \quad \text{with} \quad r \gtrsim \epsilon^{-\frac{d-1}{sd}}$$

- But low-rank approximations are expected to exploit additional features.

Approximation in low-rank tensor subsets

- Good approximation properties: for a large set of functions $C(V)$ of interest,

$$\inf_{v \in \mathcal{M}} \|u - v\| \leq \epsilon(n) \quad \forall u \in C(V)$$

  with rapidly decaying $\epsilon(n)$.
- Good approximation for smooth functions

Example (Sobolev regularity: approximation in canonical format)

$$\inf_{v \in \mathcal{R}_r} \|u - v\|_{L^2} \lesssim r^{-sd/(d-1)} \quad \forall u \in B_s^{mix} \subset L^2(\pi_d)$$

$$\text{with} \quad B_s^{mix} = \left\{ u \in L^2(\pi_d); \|u\|_{H_s^{mix}} \leq 1 \right\} \subset H_s^{mix}(\pi_d)$$

That means that for any $u \in B_s^{mix}$ and for $\epsilon > 0$, it could be possible to find an approximation $v(x_1, \ldots, x_d) = \sum_{i=1}^r \phi_i^1(x_1) \ldots \phi_i^d(x_d)$ such that

$$\|u - v\| \leq \epsilon \quad \text{with} \quad r \gtrsim \epsilon^{-\frac{d-1}{sd}}$$

- But low-rank approximations are expected to exploit additional features.

Approximation in low-rank tensor subsets

- Good approximation properties: for a large set of functions $C(V)$ of interest,

$$\inf_{v \in \mathcal{M}} \|u - v\| \leq \epsilon(n) \quad \forall u \in C(V)$$

  with rapidly decaying $\epsilon(n)$.

- Good approximation for smooth functions

Example (Sobolev regularity: approximation in canonical format)

$$\inf_{v \in \mathcal{R}_r} \|u - v\|_{L^2} \lesssim r^{-sd/(d-1)} \quad \forall u \in B_s^{mix} \subset L^2(\pi_d)$$

$$\text{with} \quad B_s^{mix} = \left\{ u \in L^2(\pi_d); \|u\|_{H_s^{mix}} \leq 1 \right\} \subset H_s^{mix}(\pi_d)$$

That means that for any $u \in B_s^{mix}$ and for $\epsilon > 0$, it could be possible to find an approximation $v(x_1, \ldots, x_d) = \sum_{i=1}^r \phi_i^1(x_1) \ldots \phi_i^d(x_d)$ such that

$$\|u - v\| \leq \epsilon \quad \text{with} \quad r \gtrsim \epsilon^{-\frac{d-1}{sd}}$$

- But low-rank approximations are expected to exploit additional features.

Best approximation in low-rank tensor subsets

- Best approximation problems in tree-based low-rank subsets $\mathcal{M}_{\leq r}$ are well-posed (provided some conditions on tensor norms).
- Best approximation problems related to singular value decompositions and their generalizations.
- For $d > 2$, no (guaranteed) algorithm for obtaining best approximations but for some tensor subsets (and particular norms), algorithms for obtaining quasi-best approximations

$$v_\gamma \in \mathcal{M} \quad \text{such that} \quad \|u - v_\gamma\| \leq (1 + \gamma(d)) \inf_{v \in \mathcal{M}} \|u - v\|$$

$$1 + \gamma(d) = \begin{cases} \sqrt{d} & \text{for Tucker tensors} \\ \sqrt{2d-2} & \text{for tree-based Tucker tensors} \end{cases}$$

(for Hilbert spaces equipped with the canonical norm)

Best approximation in low-rank tensor subsets

- Best approximation problems in tree-based low-rank subsets $\mathcal{M}_{\leq r}$ are well-posed (provided some conditions on tensor norms).
- Best approximation problems related to singular value decompositions and their generalizations.
- For $d > 2$, no (guaranteed) algorithm for obtaining best approximations but for some tensor subsets (and particular norms), algorithms for obtaining quasi-best approximations

$$v_\gamma \in \mathcal{M} \quad \text{such that} \quad \|u - v_\gamma\| \leq (1 + \gamma(d)) \inf_{v \in \mathcal{M}} \|u - v\|$$

$$1 + \gamma(d) = \begin{cases} \sqrt{d} & \text{for Tucker tensors} \\ \sqrt{2d-2} & \text{for tree-based Tucker tensors} \end{cases}$$

(for Hilbert spaces equipped with the canonical norm)

Best approximation in low-rank tensor subsets

- Best approximation problems in tree-based low-rank subsets $\mathcal{M}_{\leq r}$ are well-posed (provided some conditions on tensor norms).
- Best approximation problems related to singular value decompositions and their generalizations.
- For $d > 2$, no (guaranteed) algorithm for obtaining best approximations but for some tensor subsets (and particular norms), algorithms for obtaining quasi-best approximations

$$v_\gamma \in \mathcal{M} \quad \text{such that} \quad \|u - v_\gamma\| \leq (1 + \gamma(d)) \inf_{v \in \mathcal{M}} \|u - v\|$$

$$1 + \gamma(d) = \begin{cases} \sqrt{d} & \text{for Tucker tensors} \\ \sqrt{2d-2} & \text{for tree-based Tucker tensors} \end{cases}$$

(for Hilbert spaces equipped with the canonical norm)

Geometry of low-rank tensors

- Subsets of tensors with fixed tree-based rank have a manifold structure :

$$\mathcal{M}_{\leq r} = \bigcup_{s \leq r} \mathcal{M}_{=s}$$

$$\mathcal{M}_{=s} = \{v \in V : rank(v) = s\} = \left\{ v = F_{\mathcal{M}}(p) \;;\; p = (p_1, \ldots, p_L) \in \mathcal{P}^1 \times \ldots \times \mathcal{P}^L \right\}$$

where $F_{\mathcal{M}}$ is a multilinear map and the $\mathcal{P}^l$ are low-dimensional vector spaces (or manifolds).

- Interesting consequences:
  - Optimization algorithms on manifolds
  - Dynamical systems on low-rank manifolds

Falco, Hackbusch and Nouy.
Geometric structures in tensor representations. MIS Preprint.

Outline

Equations in tensor format

Problem to solve

$$A(u) = b \quad \text{with} \quad u \in V = \bigotimes_{\mu=1}^{d} V_\mu$$

with

$$A = \sum_i \bigotimes_{\nu=1}^{d} A_i^\nu, \quad b = \sum_i \bigotimes_{\nu=1}^{d} b_i^\nu$$

Equations in tensor format

Problem to solve

$$A(u) = b \quad \text{with} \quad u \in V = \bigotimes_{\mu=1}^{d} V_\mu$$

with

$$A = \sum_i \bigotimes_{\nu=1}^{d} A_i^\nu, \quad b = \sum_i \bigotimes_{\nu=1}^{d} b_i^\nu$$

Example (Parametric PDE)

$$-a_1(\xi)\Delta u + a_2(\xi)u + a_3(\xi)c \cdot \nabla u = g(x)\phi(\xi), \quad x \in \Omega, \quad \xi \in \Xi, \quad u = 0 \quad on \quad \partial\Omega$$

recasted under the form

$$Au = b, \quad u \in H_0^1(\Omega) \otimes L_\mu^2(\Xi)$$

with

$$A = -\Delta \otimes A_1 + I \otimes A_2 + (c \cdot \nabla) \otimes A_3, \quad A_i : \lambda(\xi) \mapsto a_i(\xi)\lambda(\xi)$$

$$b = g \otimes \phi$$

Iterative solvers and low-rank tensor approximations

Given a low-rank format $\mathcal{M}$, replace an iterative solver

$$\{u_n\}_n \subset V \quad \text{s.t.} \quad u_n = B_n(u_{n-1})$$

with

$$\{u_n\}_n \subset \mathcal{M} \quad \text{s.t.} \quad u_n = T^{\epsilon}_{\mathcal{M}} B_n(u_{n-1})$$

where $T^{\epsilon}_{\mathcal{M}}$ is a "truncation operator" such that

$$\|T^{\epsilon}_{\mathcal{M}}(v) - v\| \leq \epsilon \|v\|$$

- When $\|\cdot\|$ is a canonical inner product norm, efficient algorihms based on SVD and controlled quasi-best approximations.
- Requires robustness of iterative solvers with respect to perturbations.
- Requires good preconditioners in tensor formats

Direct low-rank approximation

Given a low-rank tensor subset $\mathcal{M}$, replace

$$\inf_{v \in \mathcal{M}} \|u - v\|$$

by the optimization of a criterium

$$\inf_{v \in \mathcal{M}} \mathcal{E}(u, v)$$

yielding a computable approximation of u in $\mathcal{M}$.

Direct low-rank approximation

Given a low-rank tensor subset $\mathcal{M}$, replace

$$\inf_{v \in \mathcal{M}} \|u - v\|$$

by the optimization of a criterium

$$\inf_{v \in \mathcal{M}} \mathcal{E}(u, v)$$

yielding a computable approximation of u in $\mathcal{M}$.

- Natural minimization problems:

$$u = \arg\min_{v \in V} \mathcal{J}(v), \quad \mathcal{E}(u, v) = \mathcal{J}(v) - \mathcal{J}(u)$$

- Residual norms:

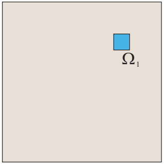
$$Au = b, \quad \mathcal{E}(u, v) = \|b - Av\|_*$$

- Sample-based semi-norms for stochastic/parametric problems:

$$\mathcal{E}(u, v) = \|u - v\|_K^2 + \text{“Regularization”} \quad \text{with} \quad \|u - v\|_K^2 = \sum_{k=1}^{Q} \|u(\xi^k) - v(\xi^k)\|^2$$

Illustration: advection-diffusion-reaction equation

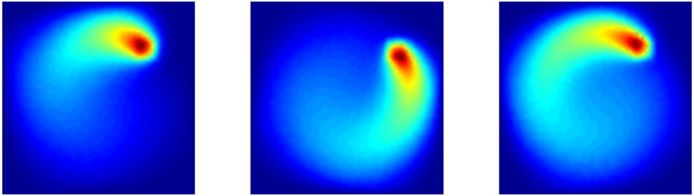
$$-a_1 \Delta u + a_2 u + a_3 c \cdot \nabla u = I_{\Omega_1} \quad on \quad \Omega$$
$$u = 0 \quad on \quad \partial\Omega$$



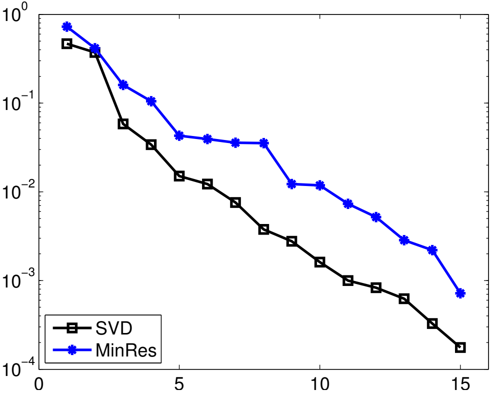
Uncertain parameters

$$a_1 = \mu_1(1 + 0.2\xi_1), \quad a_2 = \mu_2(1 + 0.2\xi_2), \quad a_3 = \xi_3$$
$$\xi_i \in U(-1, 1),\ \Xi = (-1, 1)^3$$

Samples of the solution $u(x, \boldsymbol{\xi})$



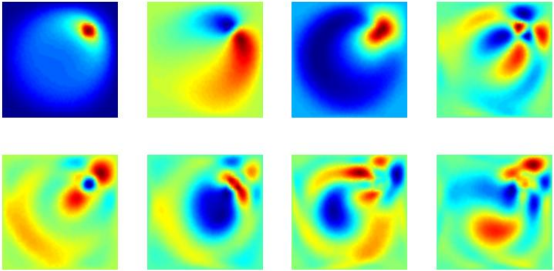
Convergence of best rank-m approximations in $L^2_\mu(\Xi;\mathcal{V})$-norm $\|\cdot\|$



$$u_m(x,\xi) = \sum_{i=1}^{m} w_i(x)\lambda_i(\xi)$$

- SVD: $u_m \in \arg\min_{v\in\mathcal{R}_m} \|u - v\|$
- MinRes: $u_m \in \arg\min_{v\in\mathcal{R}_m} \|Au - Av\|_*$ (canonical norm of the algebraic residual)

8 first spatial modes v_i of the SVD $u_m = \sum_{i=1}^{m} v_i \otimes \phi_i$



Direct approximation

If

$$\alpha\|u-v\| \leq \mathcal{E}(u,v) \leq \beta\|u-v\|$$

then

$$\tilde{u} = \arg\min_{v\in\mathcal{M}} \mathcal{E}(u,v)$$

is such that

$$\|u-\tilde{u}\| \leq \frac{\beta}{\alpha}\min_{v\in\mathcal{M}}\|u-v\|$$

Interest of working with well conditioned formulations,
i.e. such that $\beta/\alpha \approx 1$

- Construction of preconditioners in low-rank tensor formats (explicit preconditioning)

$$P = \sum_{i \in I} \alpha_i \bigotimes_{\mu=1}^{d} P^{\mu}_{i_\mu} \quad \text{(structured } \alpha\text{)}$$

[Giraldi, Nouy, Legrain 2013]

- Use non conventional minimal residual based formulations (implicit preconditioning)

$$A : V \to W', \quad \|Av - b\|_{W'} = \|v - u\|,$$

$$(1-\delta)\|u - v\| \leq \mathcal{E}(u, v) = \|\Lambda_\delta(Av - b)\|_{W'} \leq (1+\delta)\|u - v\|$$

[Cohen, Dahmen & Welper 2012]

[Billaud-Friess, Nouy & Zahm 2013]

Tensor approximation using statistical learning

- Computation of an output quantity of interest

$$f(\xi) = Q(u(\xi)), \quad \xi \sim \mu,$$

$$f \in L^2_\mu(\mathbb{R}^d) = L^2_{\mu_1}(\mathbb{R}) \otimes \ldots \otimes L^2_{\mu_d}(\mathbb{R})$$

Tensor approximation using statistical learning

- Computation of an output quantity of interest

$$f(\xi) = Q(u(\xi)), \quad \xi \sim \mu,$$

$$f \in L^2_\mu(\mathbb{R}^d) = L^2_{\mu_1}(\mathbb{R}) \otimes \ldots \otimes L^2_{\mu_d}(\mathbb{R})$$

- Best L^2 approximation in a low-rank subset $\mathcal{M}$ (not computable)

$$\min_{v \in \mathcal{M}} \|f - v\|^2 = \min_{v \in \mathcal{M}} \mathbb{E}_\mu((f(\xi) - v(\xi))^2)$$

Tensor approximation using statistical learning

- Computation of an output quantity of interest

$$f(\xi) = Q(u(\xi)), \quad \xi \sim \mu,$$

$$f \in L^2_\mu(\mathbb{R}^d) = L^2_{\mu_1}(\mathbb{R}) \otimes \ldots \otimes L^2_{\mu_d}(\mathbb{R})$$

- Best L^2 approximation in a low-rank subset $\mathcal{M}$ (not computable)

$$\min_{v \in \mathcal{M}} \|f - v\|^2 = \min_{v \in \mathcal{M}} \mathbb{E}_\mu((f(\xi) - v(\xi))^2)$$

- Low rank approximation using least-square minimization:

$$\min_{v \in \mathcal{M}} \|f - v\|_K^2$$

with $\quad \|f - v\|_K^2 = \frac{1}{K} \sum_{k=1}^{K} (f(\xi^k) - v(\xi^k))^2 \approx \mathbb{E}_\mu((f(\xi) - v(\xi))^2)$

where $\{f(\xi^k)\}_{k=1}^K$ are evaluations of f at sample points $\{\xi^k\}_{k=1}^K$ (Black-box deterministic computations).

Tensor approximation using statistical learning

- Regularization could be required

$$\min_{v \in \mathcal{M}} \|f - v\|_K^2 + \text{``regularization''}$$

Tensor approximation using statistical learning

- Regularization could be required

$$\min_{v \in \mathcal{M}} \|f - v\|_K^2 + \text{“regularization”}$$

- For a given tensor format

$$\mathcal{M} = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, 1 \leq k \leq r\}$$

solve

$$\min_{p_1, \dots p_r} \|f - F_{\mathcal{M}}(p_1, \dots, p_r)\|_K^2 + \sum_k \lambda_k \|p_k\|_s$$

that corresponds to a minimization in a subset of $\mathcal{M}$:

$$\mathcal{M}_\gamma = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, \|p_k\|_s \leq \gamma_k, 1 \leq k \leq r\}$$

Tensor approximation using statistical learning

- Regularization could be required

$$\min_{v\in\mathcal{M}} \|f - v\|_K^2 + \text{“regularization”}$$

- For a given tensor format

$$\mathcal{M} = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, 1 \le k \le r\}$$

solve

$$\min_{p_1,\dots p_r} \|f - F_{\mathcal{M}}(p_1, \dots, p_r)\|_K^2 + \sum_k \lambda_k \|p_k\|_s$$

that corresponds to a minimization in a subset of $\mathcal{M}$:

$$\mathcal{M}_\gamma = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, \|p_k\|_s \le \gamma_k, 1 \le k \le r\}$$

Tensor approximation using statistical learning

- Regularization could be required

$$\min_{v\in\mathcal{M}} \|f - v\|_K^2 + \text{``regularization''}$$

- For a given tensor format

$$\mathcal{M} = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, 1 \le k \le r\}$$

solve

$$\min_{p_1,\dots p_r} \|f - F_{\mathcal{M}}(p_1, \dots, p_r)\|_K^2 + \sum_k \lambda_k \|p_k\|_s$$

that corresponds to a minimization in a subset of $\mathcal{M}$:

$$\mathcal{M}_\gamma = \{v = F_{\mathcal{M}}(p_1, \dots, p_r); p_k \in \mathbb{R}^{m_k}, \|p_k\|_s \le \gamma_k, 1 \le k \le r\}$$

- Sparsity-inducing regularization with $0 \le s \le 1$.

Outline

Low-rank approximation and subspace-based model reduction

- Rank-m tensor subset in $\mathcal{V} \otimes \mathcal{S}$

$$\mathcal{R}_m = \{v \in \mathcal{V} \otimes \mathcal{S} \; ; \; rank(v) \leq m\} = \left\{ v = \sum_{i=1}^{m} w_i \otimes \lambda_i ; w_i \in \mathcal{V}, \lambda_i \in \mathcal{S} \right\}$$

- Subspace point of view

$$\mathcal{R}_m = \{v \in \mathcal{V}_m \otimes \mathcal{S}_m ; dim(\mathcal{V}_m) = dim(\mathcal{S}_m) = m\}$$

- Best rank-m approximation of $u \in \mathcal{V} \otimes \mathcal{S}$

$$\min_{v \in \mathcal{R}_m} \mathcal{E}(u, v) = \min_{dim(\mathcal{V}_m)=m} \; \min_{dim(\mathcal{S}_m)=m} \; \min_{v \in \mathcal{V}_m \otimes \mathcal{S}_m} \mathcal{E}(u, v)$$

This defines sequences of optimal subspaces $\mathcal{V}_m$ and $\mathcal{S}_m$.

Hilbert setting: induced norm and SVD

Let $\mathcal{V}$ and $\mathcal{S}$ be Hilbert spaces and $\|\cdot\|$ the canonical (induced) inner product norm,

$$\langle w \otimes \lambda, w' \otimes \lambda' \rangle = \langle w, w' \rangle_{\mathcal{V}} \langle \lambda, \lambda' \rangle_{\mathcal{S}}.$$

- u is identified with an operator $u : w \in \mathcal{V} \to \langle u, w \rangle_{\mathcal{V}} \in \mathcal{S}$ which is compact and admits a singular value decomposition

$$u = \sum_{i=1}^{\infty} \sigma_i w_i \otimes \lambda_i, \quad (\sigma_i) \in \ell_2(\mathbb{N})$$

- The best rank-m approximation u_m in the norm $\|\cdot\|$ coincides with the rank-m truncated singular value decomposition of u.

$$u_m = \sum_{i=1}^{m} \sigma_i w_i \otimes \lambda_i$$

- Notion of decomposition with successive optimality conditions.
- Nested subspaces $\mathcal{V}_m = span\{w_i\}_{i=1}^m$ and $\mathcal{S}_m = span\{\lambda_i\}_{i=1}^m$:

$$\mathcal{V}_m \subset \mathcal{V}_{m+1} \quad \text{and} \quad \mathcal{S}_m \subset \mathcal{S}_{m+1}$$

Low-rank approximation in Bochner Hilbert space $\mathcal{V} \otimes L^2_\mu(\Xi)$

- Natural (induced) norm

$$\|u\| = \left(\int_\Xi \|u(y)\|^2_{\mathcal{V}} \mu(dy)\right)^{1/2}$$

- A rank-m approximation u_m is of the form

$$u_m(y) = \sum_{i=1}^{m} w_i \lambda_i(y)$$

- The best rank-m approximation u_m which solves

$$\min_{v \in \mathcal{R}_m} \|u - v\|^2$$

corresponds to the truncated singular value decomposition of $U : L^2_\mu(\Xi) \to \mathcal{V}$ defined by

$$U\lambda = \int_\Xi u(y)\lambda(y)\mu(dy)$$

also known as Karhunen-Loeve decomposition (for y random variables) or Proper Orthogonal Decomposition (for y the time)

- For samples $\{y^k\}_{k=1}^K \subset \Xi$, we introduce the sample-based semi-norm

$$\|u\|_K = \left(K^{-1} \sum_{k=1}^{K} \|u(y^k)\|_{\mathcal{V}}^2 \right)^{1/2}$$

- The best rank-m approximation u_m which solves

$$\min_{v \in \mathcal{R}_m} \|u - v\|_K^2$$

corresponds to the truncated singular value decomposition of $U : \mathbb{R}^K \to \mathcal{V}$ defined by

$$U\alpha = K^{-1} \sum_{k=1}^{K} u(y^k)\alpha_k,$$

also known as Empirical Karhunen-Loeve decomposition (for y random variables) or Snapshots Proper Orthogonal Decomposition (for y the time)

Optimal low-rank approximation in the general case

In the general case (provided well-posedness of minimization problems), optimal spaces are still well defined by

$$\min_{v\in\mathcal{R}_m} \mathcal{E}(u,v) = \min_{dim(\mathcal{V}_m)=m} \min_{dim(\mathcal{S}_m)=m} \min_{v\in\mathcal{V}_m\otimes\mathcal{S}_m} \mathcal{E}(u,v)$$

Optimal low-rank approximation in the general case

In the general case (provided well-posedness of minimization problems), optimal spaces are still well defined by

$$\min_{v\in\mathcal{R}_m} \mathcal{E}(u,v) = \min_{dim(\mathcal{V}_m)=m} \min_{dim(\mathcal{S}_m)=m} \min_{v\in\mathcal{V}_m\otimes\mathcal{S}_m} \mathcal{E}(u,v)$$

BUT

- optimal subspaces are not nested

$$\mathcal{V}_m \not\subset \mathcal{V}_{m+1}, \quad \mathcal{S}_m \not\subset \mathcal{S}_{m+1}$$

- no notion of decomposition

$$u_m = \sum_{i=1}^{m} w_i^m \otimes \lambda_i^m$$

How to recover a notion of decomposition ?

Suboptimal constructions with nested subspaces and notion of decomposition

- Reduced Basis method (greedy algorithms) (for $L^\infty(\Xi) \otimes \mathcal{V}$)
- Proper Generalized Decompositions (for $L^2(\Xi) \otimes \mathcal{V}$)
- Adaptive Cross Approximation and Empirical Interpolation Method (for $L^\infty \otimes L^\infty$)

Based on greedy constructions of the approximation or of subspaces.

Best rank-m approximation in $\mathcal{V} \otimes L^\infty_\mu(\Xi)$

- Let $u \in \mathcal{V} \otimes L^\infty_\mu(\Xi)$, with $\mathcal{V}$ a Hilbert space with inner product $(\cdot,\cdot)_\mathcal{V}$ and norm $\|\cdot\|_\mathcal{V}$.

$$\|u\|_\infty = \operatorname*{ess\,sup}_{y\in\Xi} \|u(y)\|_\mathcal{V}$$

- Optimal rank-m approximation is defined by

$$\min_{v\in\mathcal{R}_m} \|u - v\|_\infty = \min_{dim(\mathcal{V}_m)} \|u - P_{\mathcal{V}_m} u\|_\infty = \min_{dim(\mathcal{V}_m)} \operatorname*{ess\,sup}_{y\in\Xi} \|u(y) - P_{\mathcal{V}_m} u(y)\|_\mathcal{V}$$

where $P_{\mathcal{V}_m} : \mathcal{V} \to \mathcal{V}_m$ is the orthogonal projector onto $\mathcal{V}_m$.

Best rank-m approximation in $\mathcal{V} \otimes L^\infty_\mu(\Xi)$

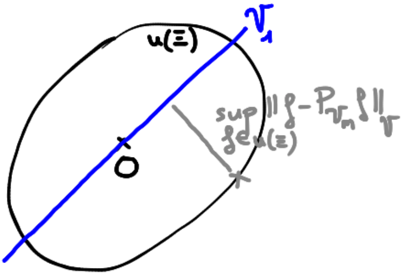
- Set of solutions

$$\mathcal{K} = \{u(y); y \in \Xi\} \subset \mathcal{V}$$

- Assuming $\mathcal{K}$ compact,

$$\min_{v \in \mathcal{R}_m} \|u - v\|_\infty = \min_{dim(\mathcal{V}_m)} \sup_{f \in \mathcal{K}} \|f - P_{\mathcal{V}_m} f\|_{\mathcal{V}} = d_m(\mathcal{K}; \mathcal{V})$$

where $d_m(\mathcal{K}; \mathcal{V})$ is the Kolmogorov m-width of the set $\mathcal{K}$.



- In general, optimal spaces are such that

$$\mathcal{V}_m \not\subset \mathcal{V}_{m+1}$$

Greedy algorithm

Suboptimal construction of subspaces $\{\mathcal{V}_m\}_{m\geq 1}$ by imposing

- $\mathcal{V}_m \subset \mathcal{V}_{m+1}$
- $\mathcal{V}_m$ has a basis composed by pointwise evaluations of $u(\xi)$.

Greedy algorithm

Start with $u_0 = 0$ and for $m \geq 1$, do

- Compute $y_m \in \Xi$ such that

$$\|u(y_m) - u_{m-1}(y_m)\|_{\mathcal{V}} = \max_{y\in\Xi} \|u(y) - u_{m-1}(y)\|_{\mathcal{V}}$$

- Let $w_m = u(y_m)$ and $\mathcal{V}_m = span\{w_1, \dots, w_m\}$
- Let u_m such that

$$u_m(y) = P_{\mathcal{V}_m} u(y)$$

Reduced Basis method using greedy algorithm

- Parametric equation

$$A(u(y); y) = b(y)$$

- Minimization of a computable surrogate error

$$\mathcal{E}(u(y), v(y); y)$$

- Introduction of a discrete parameter set $\Xi_{train} \subset \Xi$

Greedy algorithm for Reduced Basis Method

Start with $u_0 = 0$ and for $m \geq 1$, do

- Compute $y_m \in \Xi_{train}$ such that

$$\mathcal{E}(u(y_m), u_{m-1}(y_m); y_m) = \max_{y \in \Xi_{train}} \mathcal{E}(u(y), u_{m-1}(y); y)$$

- Let $w_m = u(y_m)$ and $\mathcal{V}_m = span\{w_1, \ldots, w_m\}$
- Let u_m such that

$$\mathcal{E}(u(y), u_m(y); y) = \min_{v \in \mathcal{V}_m} \mathcal{E}(u(y), v; y)$$

Greedy algorithm and Reduced Basis method

Assuming

$$\alpha\|u-v\|_{\mathcal{V}} \leq \mathcal{E}(u,v;y) \leq \beta\|u-v\|_{\mathcal{V}},$$

and $\Xi = \Xi_{train}$, then

$$\|u(y_m)-u_{m-1}(y_m)\|_{\mathcal{V}} \geq \frac{\alpha}{\beta}\max_{y\in\Xi}\|u(y)-u_{m-1}(y)\|_{\mathcal{V}}$$

A quasi-optimality result [DeVore et al. 2012]

When $\mathcal{V}$ is a Hilbert space,

$$\|u-u_m\|_{\mathcal{V}} \leq \sqrt{2}\frac{\beta}{\alpha}\min_{1\leq i<m}(d_m(\mathcal{K};\mathcal{V}))^{\frac{m-i}{m}}$$

In particular

- $\|u-u_{2m}\|_{\mathcal{V}} \leq \sqrt{2}\frac{\beta}{\alpha}\sqrt{d_m(\mathcal{K};\mathcal{V})}$
- If $d_m(\mathcal{K};\mathcal{V}) \leq C_0 m^{-\delta}$, then $\|u-u_m\| \leq C_1 m^{-\delta}$ with $C_1 = 2^{5\alpha+1}\frac{\beta^2}{\alpha^2}C_0$
- If $d_m(\mathcal{K};\mathcal{V}) \leq C_0 e^{-c_0 m^\delta}$, then $\|u-u_m\| \leq \sqrt{2C_0}\frac{\beta}{\alpha}e^{-c_1 m^\delta}$ with $c_1 = 2^{-1-2\alpha}c_0$.

- Greedy construction of the approximation (well-known version of PGD)

$$\|u - u_{m-1} - w_m \otimes \lambda_m\| = \min_{v \in \mathcal{R}_1} \|u - u_{m-1} - v\|$$

$$u_m = \sum_{i=1}^{m} w_i \otimes \lambda_i \in \mathcal{V}_m \otimes \mathcal{S}_m, \quad \mathcal{V}_m = span\{w_i\}_{i=1}^m,\ \mathcal{S}_m = span\{\lambda_i\}_{i=1}^m$$

Proper Generalized Decomposition

- Greedy construction of the approximation (well-known version of PGD)

$$\|u - u_{m-1} - w_m \otimes \lambda_m\| = \min_{v \in \mathcal{R}_1} \|u - u_{m-1} - v\|$$

$$u_m = \sum_{i=1}^{m} w_i \otimes \lambda_i \in \mathcal{V}_m \otimes \mathcal{S}_m, \quad \mathcal{V}_m = span\{w_i\}_{i=1}^m, \; \mathcal{S}_m = span\{\lambda_i\}_{i=1}^m$$

- Greedy construction of subspaces (not so well known versions of PGD !)

$$\|u - u_m\| = \min_{\substack{dim(\mathcal{V}_m)=m \\ \mathcal{V}_m \supset \mathcal{V}_{m-1}}} \min_{\substack{dim(\mathcal{S}_m)=m \\ \mathcal{S}_m \supset \mathcal{S}_{m-1}}} \min_{v \in \mathcal{V}_m \otimes \mathcal{S}_m} \|u - v\|$$

or partially greedy construction

$$\|u - u_m\| = \min_{\substack{dim(\mathcal{V}_m)=m \\ \mathcal{V}_m \supset \mathcal{V}_{m-1}}} \min_{v \in \mathcal{V}_m \otimes \mathcal{S}} \|u - v\|$$

- Suboptimal greedy construction of subspaces [N. 2008; Tamellini, Le Maitre & N. 2013, Giraldi 2012] which are very close to the construction used in Empirical Interpolation Method and Greedy algorithms for Reduces Basis methods.

- Suboptimal greedy construction of subspaces [N. 2008; Tamellini, Le Maitre & N. 2013, Giraldi 2012] which are very close to the construction used in Empirical Interpolation Method and Greedy algorithms for Reduces Basis methods.
- Suboptimal partial greedy construction of subspaces [Nouy 2007]

$$\|u - u_{m-1} - w_m \otimes \lambda_m\| = \min_{v \in \mathcal{R}_1} \|u - u_{m-1} - v\|$$

$$\|u - u_m\| = \min_{v \in \mathcal{V}_m \otimes \mathcal{S}} \|u - v\|, \quad \text{with} \quad \mathcal{V}_m = span\{w_i\}_{i=1}^m$$

$$u_m = \sum_{j=1}^{m} w_i \otimes \lambda_i^m$$

Greedy construction of a reduced basis $\{w_1, \dots, w_m, \dots\}$.

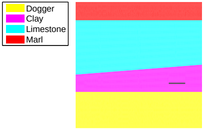
Example: stochastic Groundwater flow equation (MOMAS/Couplex)

Groundwater flow equation (hydraulic head u)

$$-\nabla(\kappa(x,\xi)\nabla u) = 0 \quad x \in \Omega,\ \xi \in \Xi$$

$+$ boundary conditions

Geological layers with uncertain properties

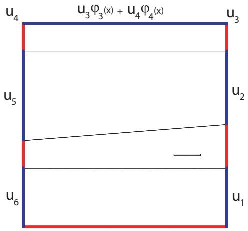


κ's probability laws

Layer	Law
Dogger	$LU(5, 125)$
Clay	$LU(3.10^{-7}, 3.10^{-5})$
Limestone	$LU(1.2, 30)$
Marl	$LU(10^{-5}, 10^{-4})$

10 basic uniform random variables ξ, $\Xi = (-1,1)^{10}$, uniform probability P_ξ

Uncertain BCs



Neumann homogeneous

Dirichlet

	Law
u_1	$U(288, 290)$
u_2	$U(305, 315)$
u_3	$U(330, 350)$
u_4	$U(170, 190)$
u_5	$U(195, 205)$
u_6	$U(285, 287)$

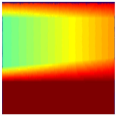
Progressive PGD

$$u_m(x,\xi) = \sum_{i=1}^{m} w_i^1(x) w_i^2(\xi)$$

Computing $(w_1^1(x), w_1^2(\xi))$ by alternated direction algorithm

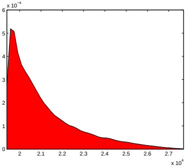
Initialize w^2 and alternate $w^1 = f_1^1(w^2)$ and $w^2 = f_1^2(w^1)$

$w^1 = f_1^1(w^2)$



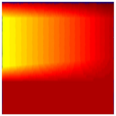
$\downarrow$

$w^2 = f_1^2(w^1)$



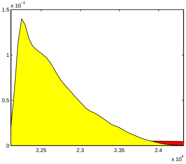
$\nearrow$

$w^1 = f_1^1(w^2)$



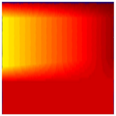
$\downarrow$

$w^2 = f_1^2(w^1)$



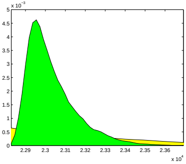
$\nearrow$

$w^1 = f_1^1(w^2)$



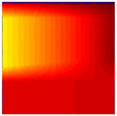
$\downarrow$

$w^2 = f_1^2(w^1)$



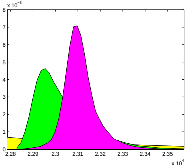
$\nearrow$

$w^1 = f_1^1(w^2)$



$\downarrow$

$w^2 = f_1^2(w^1)$

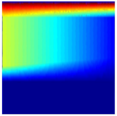


$$u_m(x,\xi) = \sum_{i=1}^{m} w_i^1(x) w_i^2(\xi)$$

Computing $(w_2^1(x), w_2^2(\xi))$ by alternated direction algorithm

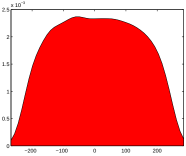
Initialize w^2 and alternate $w^1 = f_2^1(w^2)$ and $w^2 = f_2^2(w^1)$

$w^1 = f_2^1(w^2)$



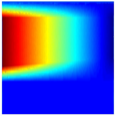
$\downarrow$

$w^2 = f_2^2(w^1)$



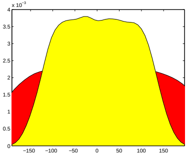
$\nearrow$

$w^1 = f_2^1(w^2)$



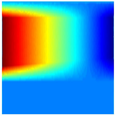
$\downarrow$

$w^2 = f_2^2(w^1)$



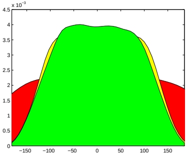
$\nearrow$

$w^1 = f_2^1(w^2)$



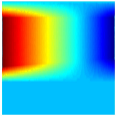
$\downarrow$

$w^2 = f_2^2(w^1)$



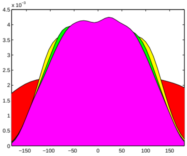
$\nearrow$

$w^1 = f_2^1(w^2)$



$\downarrow$

$w^2 = f_2^2(w^1)$

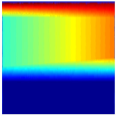


$$u_m(x,\xi) = \sum_{i=1}^{m} w_i^1(x) w_i^2(\xi)$$

Computing $(w_3^1(x), w_3^2(\xi))$ by alternated direction algorithm

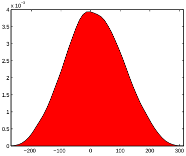
Initialize w^2 and alternate $w^1 = f_3^1(w^2)$ and $w^2 = f_3^2(w^1)$

$w^1 = f_3^1(w^2)$



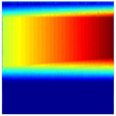
$\downarrow$

$w^2 = f_3^2(w^1)$



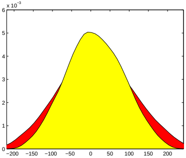
$\nearrow$

$w^1 = f_3^1(w^2)$



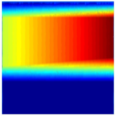
$\downarrow$

$w^2 = f_3^2(w^1)$



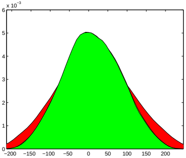
$\nearrow$

$w^1 = f_3^1(w^2)$



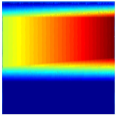
$\downarrow$

$w^2 = f_3^2(w^1)$



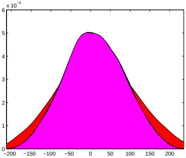
$\nearrow$

$w^1 = f_3^1(w^2)$



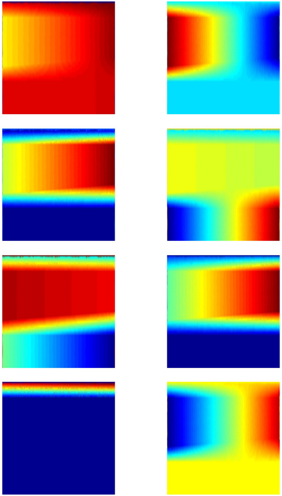
$\downarrow$

$w^2 = f_3^2(w^1)$

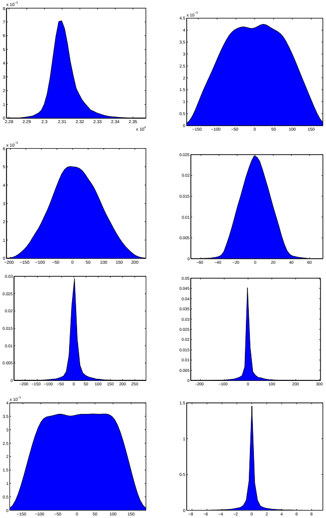


First modes of the progressive PGD

Spatial modes $\{w_1^1, \dots, w_8^1\}$

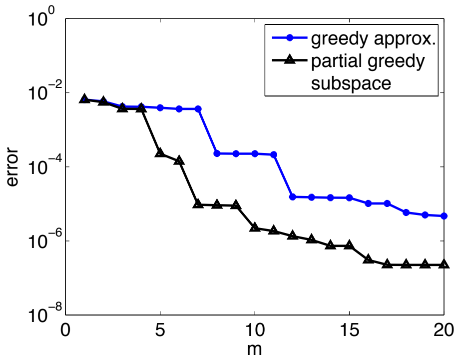


Stochastic modes $\{w_1^2, \dots, w_8^2\}$



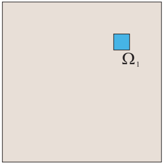
Convergence of the progressive PGD (L^2-norm)

$$\|u - u_m\|_{L^2(\Omega\times\Xi)}$$



Application to an advection-diffusion-reaction equation

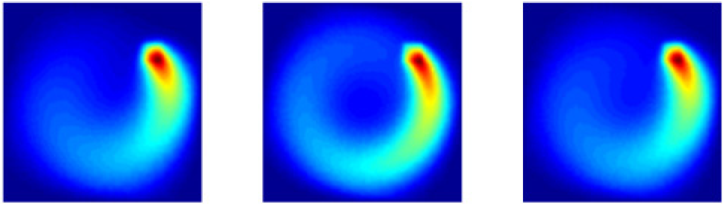
- $\partial_t u - a_1 \Delta u + a_2 c \cdot \nabla u + a_3 u = a_4 I_{\Omega_1}$ *on* $\Omega \times (0, T)$
- $u = 0$ *on* $\Omega \times \{0\}$
- $u = 0$ *on* $\partial\Omega \times (0, T)$



Uncertain parameters

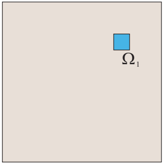
$$a_i(\boldsymbol{\xi}) = \mu_{a_i}(1 + 0.2\xi_i), \quad \xi_i \in U(-1, 1), \quad \boxed{\boldsymbol{\Xi} = (-1, 1)^4}$$

Three samples of the solution $u(x, t, \boldsymbol{\xi})$



Application to an advection-diffusion-reaction equation

- $\partial_t u - a_1 \Delta u + a_2 c \cdot \nabla u + a_3 u = a_4 I_{\Omega_1}$ *on* $\Omega \times (0, T)$
- $u = 0$ *on* $\Omega \times \{0\}$
- $u = 0$ *on* $\partial\Omega \times (0, T)$



Uncertain parameters

$a_i(\boldsymbol{\xi}) = \mu_{a_i}(1 + 0.2\xi_i), \quad \xi_i \in U(-1, 1), \quad \boxed{\boldsymbol{\Xi} = (-1, 1)^4}$

Low-rank approximation of the solution

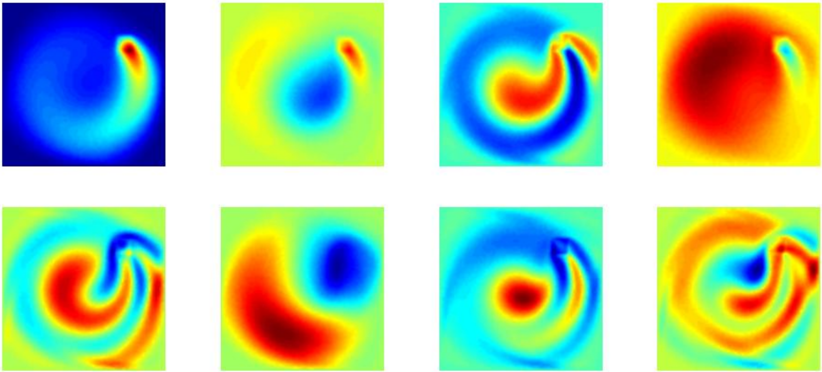
$$\boxed{u(x, t, \boldsymbol{\xi}) \approx \sum_{i=1}^{M} w_i(x, t)\lambda_i(\boldsymbol{\xi})}$$

$$w_i \in \mathcal{V} = L^2(0, T; H_0^1(\Omega)), \quad \lambda_i \in \mathcal{S} = L^2_{P_\xi}(\boldsymbol{\Xi})$$

Generalized Spectral Decomposition

Approximation of the ideal pseudo eigenspace through Arnoldi-type procedure

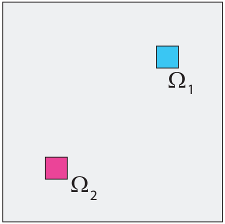
8 first modes of the decomposition $\{w_1(x,t) \ldots w_8(x,t)\}$



To compute these modes $\Rightarrow$ **only 8 deterministic problems**

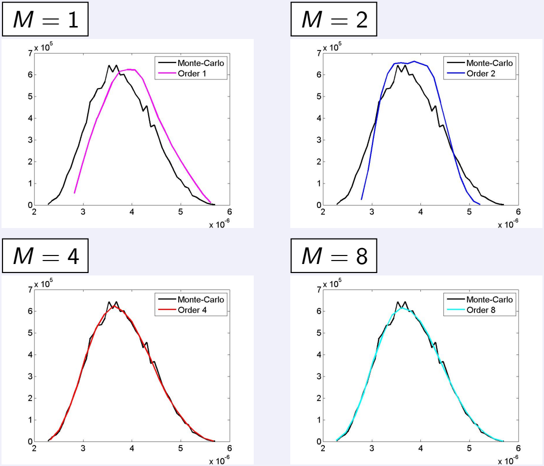
Quantity of interest

$$Q(\boldsymbol{\xi}) = \int_0^T \int_{\Omega_2} u(x, t, \boldsymbol{\xi})\, dxdt$$



$$Q_M(\boldsymbol{\xi}) = \int_0^T \int_{\Omega_2} u_M(x, t, \boldsymbol{\xi})\, dxdt$$

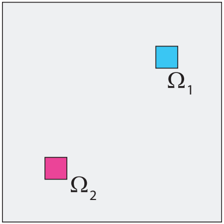
Probability density function of $Q_M(\boldsymbol{\xi})$



Convergence of quantities of interest
Quantiles

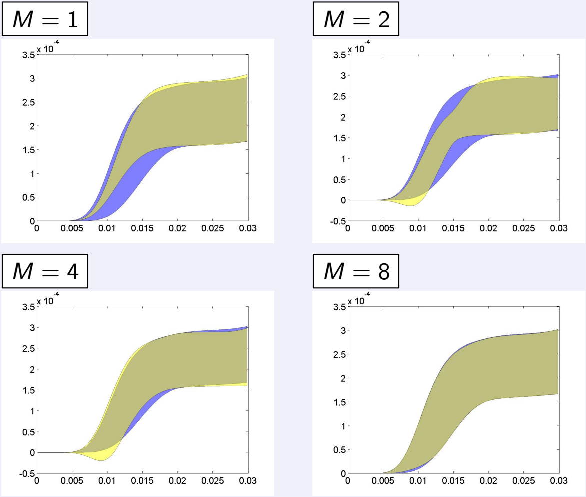
Quantity of interest

$$Q(t,\boldsymbol{\xi}) = \int_{\Omega_2} u(x,t,\boldsymbol{\xi})\,dx$$

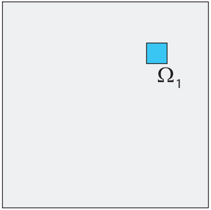


$$Q_M(t,\boldsymbol{\xi}) = \int_{\Omega_2} u_M(x,t,\boldsymbol{\xi})\,dx$$

99% Quantiles of $Q_M(t,\boldsymbol{\xi})$



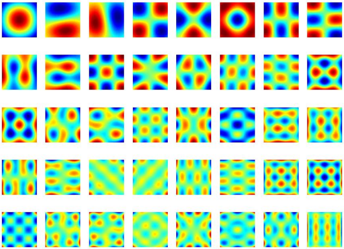
$$-\nabla \cdot (\kappa \nabla u) + c \cdot \nabla u + \gamma u = \delta I_{\Omega_1}(x) \quad on \quad \Omega$$



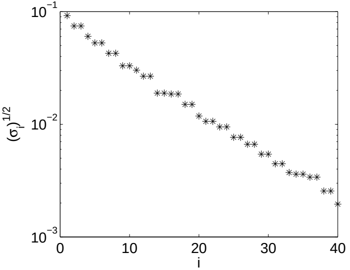
Random field

$$\kappa(x, \boldsymbol{\xi}) = \mu_\kappa + \sum_{i=1}^{40} \sqrt{\sigma_i} \kappa_i(x) \xi_i, \quad \xi_i \in U(-1, 1)$$

Spatial modes $\kappa_i(x)$



Amplitudes σ_i



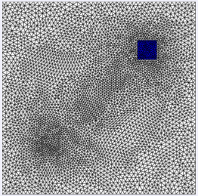
Stochastic approximation

$$\boldsymbol{\xi} = (\xi_1, \dots, \xi_{40}), \quad \boldsymbol{\Xi} = (-1, 1)^{40} = \Xi_1 \times \dots \times \Xi_{40}$$

$$\mathcal{S}_P = \mathbb{P}_4(\Xi_1) \otimes \dots \otimes \mathbb{P}_4(\Xi_{40})$$

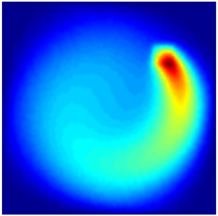
$$\boxed{dim(\mathcal{S}_P) = 5^{40} \approx 10^{28}}$$

Finite element mesh



$$\boxed{dim(\mathcal{V}_N) = 4435}$$

Solution $u(\cdot, \boldsymbol{\mu}_{\boldsymbol{\xi}})$ for mean parameters



A basic hierarchical format

Deterministic/stochastic separation

$$u(\boldsymbol{\xi}) \approx u_M(\boldsymbol{\xi}) = \sum_{i=1}^{M} w_i \lambda_i(\boldsymbol{\xi})$$

$\hookrightarrow$ $\boxed{\mathcal{V}_M = span\{w_i\}_{i=1}^M}$

Random variables separation

$$\Lambda(\boldsymbol{\xi}) := (\lambda_i)_{i=1}^M \approx \Lambda_Z(\boldsymbol{\xi}) = \sum_{k=1}^{Z} \phi_k^0 \prod_{j=1}^{s} \phi_k^j(\xi_j)$$

$\hookrightarrow$ $\boxed{\mathcal{S}_Z = span\{\prod_{j=1}^s \phi_k^j(\xi_j)\}_{k=1}^Z}$

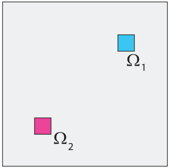
For a precision $\|u - u_{M,Z}\|_{L^2} \leqslant 10^{-2}$

- $\boxed{dim(\mathcal{V}_M) \approx 15} \ll 4435 = dim(\mathcal{V}_N)$
- $\boxed{dim(\mathcal{S}_Z) \approx 10} \ll 10^{28} = dim(\mathcal{S}_P)$
- 15 classical deterministic problems in order to build $\mathcal{V}_M \subset \mathcal{V}_N$
- about 1 minute computation on a laptop with matlab

Results

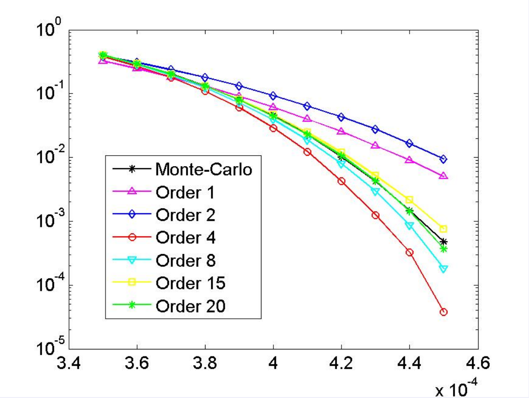
Quantity of interest

$$Q(\xi) = \int_{\Omega_2} u(x,\xi)\,dx$$



$$Q_M(\xi) = \int_{\Omega_2} u_M(x,\xi)\,dx$$

$P(Q > q), \quad q \in (3.5, 5.4)$



Convergence properties of quantities of interest

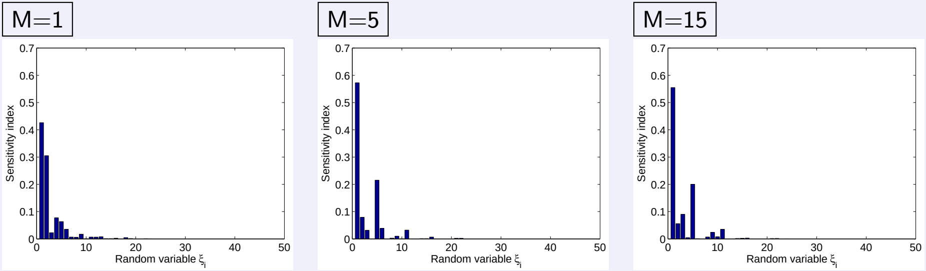
Sensitivity analysis

$$Q(\boldsymbol{\xi}) \approx Q_M(\boldsymbol{\xi}) \approx Q_{M,Z}(\boldsymbol{\xi}) = \sum_{k=1}^{Z} q_k \Psi_k(\boldsymbol{\xi}), \quad \Psi_k(\boldsymbol{\xi}) = \prod_{i=1}^{40} \phi_k^i(\xi_i)$$

First order Sobol sensitivity index with respect to parameter ξ_i

$$\boxed{S_i = \frac{Var(E(Q|\xi_i))}{Var(Q)}} \quad E(Q|\xi_i) = \sum_{k=1}^{Z} \alpha_k^i \phi_k^i(\xi_i), \quad \alpha_k^i = q_k \prod_{\substack{j=1 \\ j \neq i}}^{40} E(\phi_k^j(\xi_j))$$

First order Sobol sensivity indices S_i



Outline

- Tucker tensors with bounded multilinear rank:

$$\mathcal{T}_r = \{v : rank_\mu(v) \leq r_\mu, \forall \mu\} = \left\{ v \in \bigotimes_{\mu=1}^{d} U_\mu \; : \; dim(U_\mu) \leq r_\mu, \forall \mu \right\}$$

- Best approximation — a subspace point of view:

$$\min_{v \in \mathcal{T}_r} \mathcal{E}(u,v) = \min_{dim(U_1)=r_1} \dots \min_{dim(U_d)=r_d} \min_{v \in U_1 \otimes \dots \otimes U_d} \mathcal{E}(u,v)$$

This yields sequences of optimal but non necessarily nested subspaces $\{U_\mu^{r_\mu} : r_\mu \geq 1\}$.

Proper Generalized Decomposition for higher-order tensors

- Greedy construction of subspaces with nestedness property

$$\mathcal{E}(u, u_m) = \min_{U_1^m \supset U_1^{m-1}} \dots \min_{U_d^m \supset U_d^{m-1}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Suboptimal greedy construction of subspaces with nestedness property (isotropic enrichment) [Giraldi, Legrain and N. 2013]

$$\mathcal{E}(u, u_{m-1} + \otimes_{\mu=1}^d w_m^{(\mu)}) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w), \quad U_\mu^m = U_\mu^{m-1} + span\{w_m^{(\mu)}\}$$

$$\mathcal{E}(u, u_m) \leq (1 + \epsilon) \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Greedy construction of subspaces with nestedness property (anisotropic enrichment) At iteration m, select dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

Proper Generalized Decomposition for higher-order tensors

- Greedy construction of subspaces with nestedness property

$$\mathcal{E}(u, u_m) = \min_{U_1^m \supset U_1^{m-1}} \dots \min_{U_d^m \supset U_d^{m-1}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Suboptimal greedy construction of subspaces with nestedness property (isotropic enrichment) [Giraldi, Legrain and N. 2013]

$$\mathcal{E}(u, u_{m-1} + \otimes_{\mu=1}^d w_m^{(\mu)}) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w), \quad U_\mu^m = U_\mu^{m-1} + span\{w_m^{(\mu)}\}$$

$$\mathcal{E}(u, u_m) \leq (1 + \epsilon) \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Greedy construction of subspaces with nestedness property (anisotropic enrichment) At iteration m, select dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

Proper Generalized Decomposition for higher-order tensors

- Greedy construction of subspaces with nestedness property

$$\mathcal{E}(u, u_m) = \min_{U_1^m \supset U_1^{m-1}} \dots \min_{U_d^m \supset U_d^{m-1}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Suboptimal greedy construction of subspaces with nestedness property (isotropic enrichment) [Giraldi, Legrain and N. 2013]

$$\mathcal{E}(u, u_{m-1} + \otimes_{\mu=1}^d w_m^{(\mu)}) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w), \quad U_\mu^m = U_\mu^{m-1} + span\{w_m^{(\mu)}\}$$

$$\mathcal{E}(u, u_m) \leq (1+\epsilon) \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, select dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

Simple Benchmark: Poisson equation

$$-\Delta u = 1 \quad \text{on} \quad \Omega = (0,1)^d$$

$$\mathcal{E}(v,u) = \mathcal{J}(v) - \mathcal{J}(u), \quad \mathcal{J}(v) = \int_\Omega \nabla v \cdot \nabla v - 2\int_\Omega v$$

$$\mathcal{E}(v,u) = \|v-u\|^2_{H^1_0}$$

- Tensor-structured operator

$$\Delta = \partial_1^2 \otimes I \otimes \ldots \otimes I + \ldots + I \otimes \ldots \otimes I \otimes \partial_d^2$$

- Finite element space $V = V_1 \otimes \ldots \otimes V_d \subset H^1_0(\Omega)$.
- Galerkin approximation $u \in V$: tensor structured algebraic equation

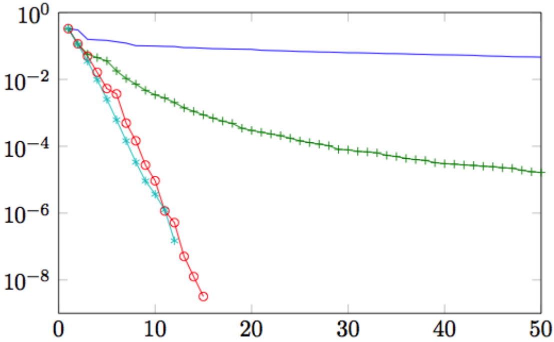
$$Au = b,$$

$$A = K \otimes M \otimes \ldots \otimes M + \ldots + M \otimes \ldots \otimes M \otimes K,$$

$$b = a \otimes \ldots \otimes a$$

Simple Benchmark: Poisson equation in dimension $d = 8$

Error with respect to the rank for different tensor formats and algorithms



$+$ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

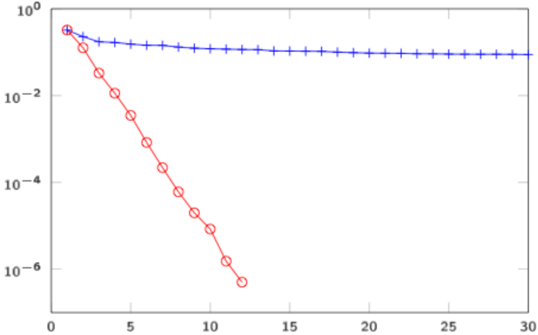
$+$ "Optimal approximation" $\mathcal{R}_r$

$\star$ Optimal approximation in $\mathcal{H}_r(V_1 \otimes \ldots \otimes V_d)$

$\circ$ Suboptimal greedy construction of subspaces U_μ^r (approximation in $\mathcal{H}_r(U_1^r \otimes \ldots \otimes U_d^r)$)

Simple Benchmark: Poisson equation in dimension $d = 27$

Error with respect to the rank for different tensor formats and algorithms



$+$ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

$\circ$ Suboptimal greedy construction of subspaces U_μ^r (approximation in $\mathcal{H}_r(U_1^r \otimes \ldots \otimes U_d^r)$)

Illustration: PDE with random coefficients

$$-\nabla \cdot (K \nabla u) + \xi_2 u = 1 \quad on \quad \Omega, \quad u = 0 \quad on \quad \partial\Omega$$

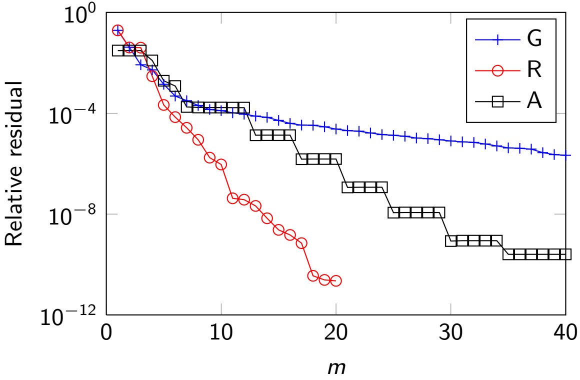
$$K = 1 + \xi_1 I_D(x)$$

$$\xi_1 \sim U(0, 10), \quad \xi_2 \sim U(0, 1)$$

$$u \in H_0^1(\Omega) \otimes L^2(\Xi_1) \otimes L^2(\Xi_2)$$

Illustration: PDE with random coefficients

Error with respect to iteration for different tensor formats and algorithms

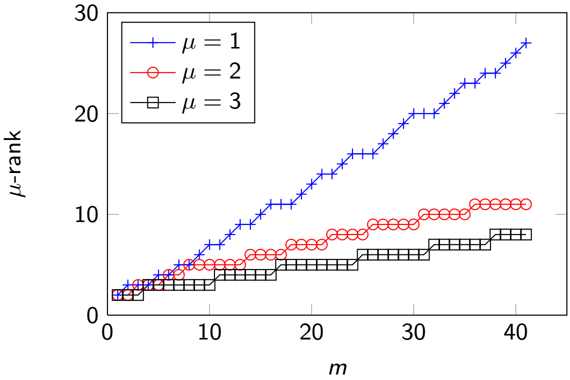


+ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

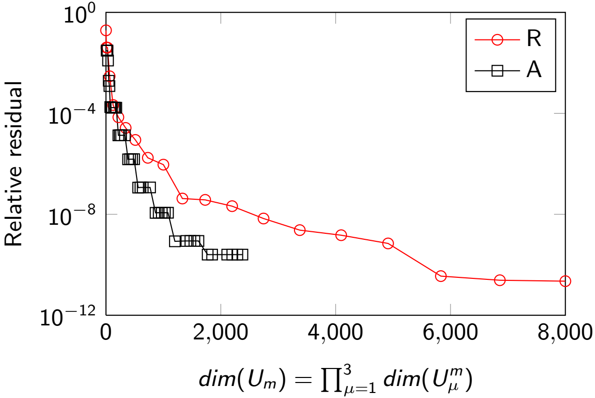
○ Suboptimal greedy construction of subspaces based on rank-one corrections (isotropic)

□ Suboptimal greedy construction of subspaces with anisotropic enrichment

Ranks with respect to iteration m for anisotropic construction



Error with respect to the dimension of the reduced space U^m for greedy constructions of subspaces.



∘ Suboptimal greedy construction of subspaces based on rank-one corrections (isotropic)

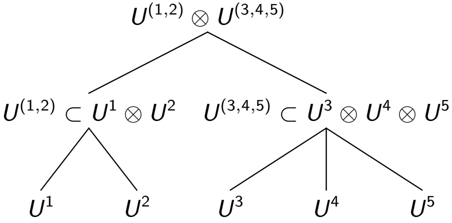
□ Suboptimal greedy construction of subspaces with anisotropic enrichment

Proper Generalized Decompositions for tree-based formats

- Tensors with bounded tree-based rank

$$\mathcal{H}_r^T = \{v \ : \ v \in U_\alpha \otimes U_{\alpha^c}, dim(U_\alpha) \leq r_\alpha, \alpha \in T\}$$

s.t. the set of subspaces $\{U_\alpha\}_{\alpha \in T}$ has a hierarchical structure



- Best approximation problems - a subspace point of view

$$\min_{v \in \mathcal{H}_r^T} \|u - v\| = \min_{\substack{U_\mu \subset V_\mu \\ dim(U_\mu) = r_\mu \\ \mu \in \{1,\dots d\}}} \ \min_{\substack{U_\alpha \subset \bigotimes_{\beta \in S(\alpha)} U_\beta \\ dim(U_\alpha) = r_\alpha \\ \alpha \in I(T)}} \ \min_{v \in \bigotimes_{\beta \in S(D)} U_\beta} \|u - v\|$$

define sequences of optimal and non necessarily nested subspaces $\{U_\alpha^{r_\alpha}; r_\alpha \geq 1\}$.

- Algorithms for the construction of suboptimal sequences of nested subspaces ... strategies of enrichment for non isotropic constructions ?

Challenging issues

- Classify applications and dedicated reduced order methods
  - Quantum physics : a long history for the construction of tensor formats
  - Machine learning and statistical learning: a huge literature on reduced order models for high dimensional functions.
- A priori estimates
- Automatic selection of reduced order formats (bases or frames for sparse approximation, tensor formats for low-rank approximation).
- Well-conditioned formulations for (quasi-)optimal model reduction
- Samples-based constructions: How to sample given an approximation format ? How many samples ?
- Software engineering. Minimize interactions with existing codes.
- Goal-oriented model order reduction: quantity of interest $Q(u)$, rare event computation, sensitivity analysis, optimization, inverse problems, ...

References I

- **LOW-RANK TENSOR METHODS**

W. Hackbusch. *Tensor Spaces and Numerical Tensor Calculus*, Springer, 2012.

L. Grasedyck and D. Kressner and C. Tobler. A literature survey of low-rank tensor approximation techniques. 2013

F. Chinesta, P. Ladeveze, and E. Cueto. A short review on model order reduction based on proper generalized decomposition. *Archives of Computational Methods in Engineering*, 18(4):395–404, 2011.

A. Falco, W. Hackbusch and A. Nouy. Geometric structure in tensor representations. Preprint.

A. Falco and A. Nouy. Proper generalized decomposition for nonlinear convex problems in tensor Banach spaces. Numerische Mathematik, 2012.

L. Giraldi, A. Nouy and G. Legrain. Low-rank approximate inverse for preconditioning tensor-structured linear systems. Preprint.

M. Billaud-Friess, A. Nouy and O. Zahm. A tensor approximation method based on ideal minimal residual formulations for the solution of high-dimensional problems. Preprint.

- **LOW-RANK FOR UQ**

Boris N. Khoromskij and Christoph Schwab.
Tensor-structured galerkin approximation of parametric and stochastic elliptic pdes.
SIAM Journal on Scientific Computing, 33(1):364–385, 2011.

References II

- A. Nouy. Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems. Archives Comp. Meth. Eng., 2010.
- M. Chevreuil, A. Nouy, P. Rai. A least-squares method for sparse low rank approximation of multivariate functions. Preprint
- L. Tamellini, O. Le Maitre, and A. Nouy. Model reduction based on proper generalized decomposition for the stochastic steady incompressible navier-stokes equations. *MOX–Report No. 26/2012*, 2012.

- **SPARSE METHODS FOR UQ**

- A. Cohen, A. Chkifa, and C. Schwab. Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs.
- A. Chkifa, A. Cohen, R. DeVore, and C. Schwab. Sparse adaptive taylor approximation algorithms for parametric and stochastic elliptic pdes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47:253–280, 1 2013.
- F. Nobile, R. Tempone, and C. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.
- I. Babuska, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Review*, 52(2):317–355, 2010.
- A. Doostan and H. Owhadi. A non-adapted sparse approximation of pdes with stochastic inputs. *Journal of Computational Physics*, 230(8):3015–3034, 2011.

References III

- **REDUCED BASIS**

A.T. Patera and G. Rozza. *Reduced Basis Approximation and A-Posteriori Error Estimation for Parametrized PDEs*. MIT-Pappalardo Graduate Monographs in Mechanical Engineering, 2007.

S. Boyaval, C. Le Bris, T. Lelievre, Y. Maday, N. C. Nguyen, and A. T. Patera. Reduced Basis Techniques for Stochastic Problems. *Archives Of Computational Methods In Engineering*, 17(4):435–454, 2010.

Y. Maday, N. C. Nguyen, A. T. Patera, and G. S. H. Pau. A general multipurpose interpolation procedure: the magic points. *Communications On Pure and Applied Analysis*, 8(1):383–404, 2009.

R. DeVore, G. Petrova, and P. Wojtaszczyk. Greedy algorithms for reduced bases in banach spaces. *Constructive Approximation*, 37(3):455–466, 2013.