Low-rank methods and Proper Generalized Decompositions — Application to parametric and stochastic problems

Anthony Nouy

Ecole Centrale Nantes / GeM
Nantes, France

Stochastic and parametric analyses

Stochastic/parametric model

$$u : \Xi \rightarrow \mathcal{V} \quad \text{such that} \quad \mathcal{F}(u(\xi); \xi) = 0$$

where ξ are parameters or random variables taking values in a measure space (Ξ, μ).

- Forward problem:
  Given μ, compute a variable of interest
  $$s(\xi) = \ell(u(\xi); \xi)$$
  and quantities of interest (e.g. statistical moments, probability of events, sensitivity indices)
- Optimization or inverse problem:
  Given observations of $s(\xi)$, determine ξ or estimate μ.

About notations

- $\Xi \subset \mathbb{R}^d$: parameter set
- μ : finite measure on Ξ
- ξ : parameters values or Ξ-valued random variable with probability law μ.
- $\mathcal{V}$: solution space for the physical model

Ideal approach

Compute an accurate approximation of $u(\xi)$ (reduced order model, metamodel, surrogate...) that allows fast evaluations of output variables of interest, observables, or objective function.

Uncertainty quantification using functional approaches

Ideal approach

Compute an accurate approximation of $u(\xi)$ (reduced order model, metamodel, surrogate...) that allows fast evaluations of output variables of interest, observables, or objective function.

Issues

1. Complex numerical models

$$u(\xi) \in \mathcal{V}_{\mathcal{N}}, \quad \mathcal{F}_{\mathcal{N}}(u(\xi);\xi) = 0$$

$$\mathcal{N} \gg 1$$

2. Approximation of multivariate functions

$$u(\xi_1,\ldots,\xi_d),\ (\xi_1,\ldots,\xi_d) \in \Xi \subset \mathbb{R}^d$$

$$d \gg 1$$

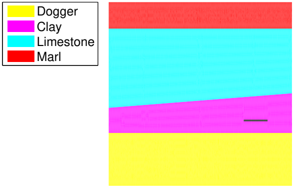
Example: Diffusion in heterogeneous media

Diffusion equation with a random diffusion coefficient κ:

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

- Groundwater flow. Geological layers with uncertain properties:

$$\kappa(x, \omega) = \sum_{i=1}^{4} \xi_i(\omega) I_{D_i}(x)$$

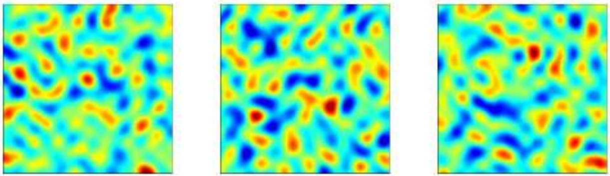


Layer	Probability Law
D_1 : Dogger	$\xi_1 \sim LU(5, 125)$
D_2 : Clay	$\xi_2 \sim LU(3.10^{-7}, 3.10^{-5})$
D_3 : Limestone	$\xi_3 \sim LU(1.2, 30)$
D_4 : Marl	$\xi_4 \sim LU(10^{-5}, 10^{-4})$

$$u(\xi) \in \mathcal{V}_{\mathcal{N}}, \quad \mathcal{N} \gg 1$$

- Random media with spatially correlated random fields

$$\kappa(x,\omega) = \exp(\underline{g}(x) + \sum_{i=1}^{d} \sqrt{\sigma_i} g_i(x) \xi_i(\omega)), \quad d \gg 1$$



Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

High dimensional problems

- Many problems in computational science and engineering require the approximation of multivariate functions:

$$u(x_1, \dots, x_d)$$

- Classical discretization methods introduce high-dimensional parametrizations

$$u(x_1, \dots, x_d) \approx \sum_{i_1=1}^{n_1} \dots \sum_{i_d=1}^{n_d} a_{i_1 \dots i_d} \varphi^1_{i_1}(x_1) \dots \varphi^1_{i_d}(x_d)$$

$$a \in \mathbb{R}^{n_1 \times \dots \times n_d}$$

- 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

$$u(x_1^k, \dots, x_d^k)$$

  - Equations (ODE, PDE, DAE)
  - Partial pointwise evaluations (e.g. for parametric/stochastic problems): equations for

$$u(\cdot, x_2^k, \dots, x_d^k)$$

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
    $$u(\cdot, x_2^k, \ldots, x_d^k)$$

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Low rank tensor approximation :

What is a tensor ?

What is a tensor-structured problem ?

What is the rank of a tensor ?

Tensor spaces

An algebraic tensor space $V = V_1 \otimes \ldots \otimes V_d$ 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 $\|\cdot\|$:

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

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

Examples of (Banach) tensor spaces

- Multidimensional array

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

$$u = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} u_{i_1,\ldots,i_d} e^1_{i_1} \otimes \ldots \otimes e^d_{i_d}$$

- Finite dimensional tensor spaces:

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

Denoting $\{\phi^k_i\}_{i=1}^{n_k}$ a basis of the n_k-dimensional space V_k, $u \in V$ can be written

$$u = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} a_{i_1 \ldots i_d} \phi^1_{i_1} \otimes \ldots \otimes \phi^d_{i_d},$$

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

- Multidimensional array

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

$$u = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} u_{i_1,\ldots,i_d} e^1_{i_1} \otimes \ldots \otimes e^d_{i_d}$$

- Finite dimensional tensor spaces:

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

Denoting $\{\phi^k_i\}_{i=1}^{n_k}$ a basis of the n_k-dimensional space V_k, $u \in V$ can be written

$$u = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} a_{i_1 \ldots i_d} \phi^1_{i_1} \otimes \ldots \otimes \phi^d_{i_d},$$

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

$$\|u\|_p = \left(\int_\Xi \|u(\xi)\|^p_\mathcal{V}\mu(d\xi)\right)^{1/p} \qquad (1\le p<\infty),$$

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

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

$$u(\xi) = \sum_{i=1}^m v_i s_i(\xi), \quad \xi\in\Xi.$$

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

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

- Case $p=\infty$.

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

Examples of (Banach) tensor spaces

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

$$\|u\|_p = \left(\int_\Xi \|u(\xi)\|_\mathcal{V}^p \mu(d\xi)\right)^{1/p} \quad (1\leq p<\infty),$$

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

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

$$u(\xi) = \sum_{i=1}^m v_i s_i(\xi), \quad \xi\in\Xi.$$

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

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

- Case $p=\infty$.

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

Examples of (Banach) tensor spaces

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

$$\|u\|_p = \left(\int_\Xi \|u(\xi)\|_\mathcal{V}^p \mu(d\xi)\right)^{1/p} \qquad (1 \le p < \infty),$$

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

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

$$u(\xi) = \sum_{i=1}^m v_i s_i(\xi), \quad \xi \in \Xi.$$

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

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

  - Case $p=\infty$.

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

Examples of (Banach) tensor spaces

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

$$\|u\|_p = \left(\int_\Xi \|u(\xi)\|^p_\mathcal{V}\mu(d\xi)\right)^{1/p} \qquad (1\leq p<\infty),$$

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

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

$$u(\xi)=\sum_{i=1}^m v_i s_i(\xi), \quad \xi\in\Xi.$$

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

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

  - Case $p=\infty$.

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

Examples of (Banach) tensor spaces

- Lebesgue space $L^p_\mu(\Xi)$ with product measure $\mu = \mu_1 \otimes \ldots \otimes \mu_d$ on $\Xi = \Xi_1 \times \ldots \times \Xi_d$:

$$L^p_\mu(\Xi_1 \times \ldots \times \Xi_d) = \overline{L^p_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^p_{\mu_1}(\Xi_d)}^{\|\cdot\|_p} \quad (1 \leq p < \infty)$$

An element $u \in L^p_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^p_{\mu_1}(\Xi_d)$ is of the form

$$u(\xi_1, \ldots, \xi_d) = \sum_{i=1}^m u_i^1(\xi_1) \ldots u_i^d(\xi_d), \quad (\xi_1, \ldots, \xi_d) \in \Xi.$$

Examples of (Banach) tensor spaces

- Lebesgue space $L^p_\mu(\Xi)$ with product measure $\mu = \mu_1 \otimes \ldots \otimes \mu_d$ on $\Xi = \Xi_1 \times \ldots \times \Xi_d$:

$$L^p_\mu(\Xi_1 \times \ldots \times \Xi_d) = \overline{L^p_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^p_{\mu_1}(\Xi_d)}^{\|\cdot\|_p} \quad (1 \leq p < \infty)$$

An element $u \in L^p_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^p_{\mu_1}(\Xi_d)$ is of the form

$$u(\xi_1, \ldots, \xi_d) = \sum_{i=1}^m u_i^1(\xi_1) \ldots u_i^d(\xi_d), \quad (\xi_1, \ldots, \xi_d) \in \Xi.$$

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

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

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

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

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

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

Tensor-structured problems in Uncertainty Quantification

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

$$\mathcal{F}(u(\xi);\xi) = 0, \quad u(\xi) \in \mathcal{V}$$

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

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

- Functions of independent random variables:

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

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

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

Tensor-structured problems in Uncertainty Quantification

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

$$\mathcal{F}(u(\xi); \xi) = 0, \quad u(\xi) \in \mathcal{V}$$

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

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

- Functions of independent random variables:

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

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

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

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

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

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

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

- 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)

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

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

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

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

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

- 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)

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

Other tensor structured problems in computational science

- Dynamical systems:

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

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

$$u \in \overline{H^1(I) \otimes \mathcal{V}}$$

- Multidimensional PDEs

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

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

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

- 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:

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

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

$$u \in \overline{H^1(I) \otimes \mathcal{V}}$$

- Multidimensional PDEs

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

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

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

- 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:

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

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

$$u \in \overline{H^1(I) \otimes \mathcal{V}}$$

- Multidimensional PDEs

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

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

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

- 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:

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

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

$$u \in \overline{H^1(I) \otimes \mathcal{V}}$$

- Multidimensional PDEs

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

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

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

- 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.

$$v \in V_1 \otimes V_2$$

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \quad \left(v(x_1, x_2) = \sum_{i=1}^{r} v_i^1(x_1) v_i^2(x_2) \right)$$

- 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.

$$v \in V_1 \otimes V_2$$

$$rank(v) \leq r \quad \Longleftrightarrow \quad v = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \quad \left(v(x_1, x_2) = \sum_{i=1}^{r} v_i^1(x_1) v_i^2(x_2) \right)$$

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

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

A class of parametric and stochastic models

$$u(\xi) \in \mathcal{V}, \quad a(u(\xi), w; \xi) = f(w; \xi) \quad \forall w \in \mathcal{W}$$

- Assumption on the parametrized bilinear form $a(\cdot, \cdot; \xi) : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$

$$\sup_{v \in \mathcal{V}} \sup_{w \in \mathcal{W}} \frac{a(v, w; \xi)}{\|v\|_{\mathcal{V}} \|w\|_{\mathcal{W}}} \leq \gamma(\xi) \leq \overline{\gamma} < \infty$$

$$\inf_{v \in \mathcal{V}} \sup_{w \in \mathcal{W}} \frac{a(v, w; \xi)}{\|v\|_{\mathcal{V}} \|w\|_{\mathcal{W}}} \geq \alpha(\xi) \geq \underline{\alpha} > 0$$

Example 1

$$-\nabla \cdot (\kappa(\cdot, \xi) \nabla u) = g(\cdot, \xi) \quad \text{on} \quad D, \quad u = 0 \quad \text{on} \quad \partial D$$

- $a(u, w; \xi) = \int_D \nabla w(x) \cdot \kappa(x, \xi) \cdot \nabla u(x) \, dx, \quad f(w; \xi) = \int_D g(x, \xi) w(x) \, dx$
- Approximation space $\mathcal{V} \subset H^1_0(D)$, $\mathcal{W} = \mathcal{V}$.
- $\underline{\alpha} \leq \alpha(\xi) \leq \kappa(x, \xi) \leq \gamma(\xi) \leq \overline{\gamma}$

- Corresponding operator equation

$$A(\xi)u(\xi) = f(\xi)$$

$$A(\xi) : \mathcal{V} \to \mathcal{W}' \quad \text{such that} \quad a(v, w; \xi) = \langle A(\xi)v, w \rangle$$

$$f(\xi) \in \mathcal{W}' \quad \text{such that} \quad f(w; \xi) = \langle f(\xi), w \rangle$$

- Operator $A(\xi)$ is such that

$$\alpha(\xi)\|v\|_{\mathcal{V}} \leq \|A(\xi)v\|_{\mathcal{W}'} \leq \gamma(\xi)\|v\|_{\mathcal{V}}$$

- Given bases $\{\varphi_i\}_{i=1}^{\mathcal{N}}$ and $\{\psi_i\}_{i=1}^{\mathcal{N}}$ of $\mathcal{V}$ and $\mathcal{W}$, algebraic formulation

$$\mathbf{u}(\xi) \in \mathbb{R}^{\mathcal{N}}, \quad \mathbf{A}(\xi)\mathbf{u}(\xi) = \mathbf{f}(\xi)$$

with $(\mathbf{A}(\xi))_{ij} = \langle A\varphi_j, \psi_i \rangle$, $(\mathbf{f}(\xi))_i = \langle f(\xi), \psi_i \rangle$, and $u(\xi) = \sum_{j=1}^{\mathcal{N}} (\mathbf{u}(\xi))_j \varphi_j$.

- Regularity of the solution

$$\|u(\xi)\|_{\mathcal{V}} \leq \frac{1}{\alpha(\xi)}\|f(\xi)\|_{\mathcal{W}'}$$

If $\alpha(\xi) \geq \underline{\alpha}$ and $f \in L^p_\mu(\Xi;\mathcal{W}')$, then

$$\|u\|_p = \mathbb{E}_\mu(\|u(\xi)\|^p_{\mathcal{V}})^{1/p} \leq \mathbb{E}_\mu(\frac{1}{\alpha(\xi)}\|f(\xi)\|^p_{\mathcal{W}'})^{1/p} \leq \frac{1}{\underline{\alpha}}\|f\|_p$$

which implies

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

- From now on, assume that

$$u \in L^2_\mu(\Xi;\mathcal{V}) = \overline{\mathcal{V} \otimes L^2_\mu(\Xi)}^{\|\cdot\|_2}$$

Stochastic (or parametric) weak form

Galerkin approximation of the solution in $\overline{\mathcal{V} \otimes L^2_\mu(\Xi)}^{\|\cdot\|_2}$ defined by

$$u \in \mathcal{V} \otimes \mathcal{S}, \quad B(u,w) = F(w) \quad \forall w \in \mathcal{W} \otimes \widetilde{\mathcal{S}}$$

- Approximation spaces $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ in $L^2_\mu(\Xi)$ (e.g. polynomial chaos). Usually, $\mathcal{S} = \widetilde{\mathcal{S}}$ (Parametric Bubnov-Galerkin).
- $B(v,w) = \mathbb{E}_\mu(\langle A(\xi)v(\xi), w(\xi)\rangle) = \int_\Xi \langle A(y)v(y), w(y)\rangle \mu(dy)$
- $F(w) = \mathbb{E}_\mu(\langle f(\xi), w(\xi)\rangle) = \int_\Xi \langle f(y), w(y)\rangle \mu(dy)$
- Corresponding operator equation:

$$Bu = F$$

  with $B : \mathcal{V} \otimes \mathcal{S} \to (\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ and $F \in (\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ defined by

$$\langle Bu, w\rangle = B(u,w), \quad F(w) = \langle F, w\rangle$$

Tensor structured equations

- Low-rank representations of operator and right-hand side

$$a(v,w;\xi)=\sum_{k=1}^{R}\lambda_k(\xi)a_k(v,w),\quad A(\xi)=\sum_{k=1}^{R}\lambda_k(\xi)A_k$$

$$f(\xi)=\sum_{k=1}^{L}\eta_k(\xi)f_k$$

Example 1

- $\kappa(x,\xi)=\sum_{k=1}^{R}\lambda_k(\xi)\kappa_k(x),\quad a_k(v,w)=\int_D \nabla w(x)\cdot\kappa_k(x)\cdot\nabla v(x)\,dx=\langle A_k v,w\rangle$
- $g(\cdot,\xi)=\sum_{k=1}^{L}\eta_k(\xi)g_k(x),\quad \langle f_k,w\rangle=\int_D g_k(x)w(x)\,dx$
- If κ and g are not of this form (or if R and L are high), low-rank approximation (e.g. using SVD or Empirical Interpolation method).

Tensor structured equations

- $\lambda : \Xi \to \mathbb{R}$ can be identified with an operator $\Lambda : \mathcal{S} \to \widetilde{\mathcal{S}}'$ such that

$$\langle \Lambda s, \tilde{s} \rangle = \mathbb{E}_\mu(\lambda(\xi) s(\xi) \tilde{s}(\xi))$$

Tensor structured equations

- $\lambda : \Xi \to \mathbb{R}$ can be identified with an operator $\Lambda : \mathcal{S} \to \widetilde{\mathcal{S}}'$ such that

$$\langle \Lambda s, \tilde{s} \rangle = \mathbb{E}_\mu(\lambda(\xi) s(\xi) \tilde{s}(\xi))$$

- $A(\xi) = \sum_{k=1}^{R} \lambda_k(\xi) A_k$ defines an operator B from $\mathcal{V} \otimes \mathcal{S}$ to $(\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ such that

$$B = \sum_{k=1}^{R} A_k \otimes \Lambda_k$$

Tensor structured equations

- $\lambda : \Xi \to \mathbb{R}$ can be identified with an operator $\Lambda : \mathcal{S} \to \widetilde{\mathcal{S}}'$ such that

$$\langle \Lambda s, \tilde{s} \rangle = \mathbb{E}_\mu(\lambda(\xi) s(\xi) \tilde{s}(\xi))$$

- $A(\xi) = \sum_{k=1}^{R} \lambda_k(\xi) A_k$ defines an operator B from $\mathcal{V} \otimes \mathcal{S}$ to $(\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ such that

$$B = \sum_{k=1}^{R} A_k \otimes \Lambda_k$$

- $f(\xi) = \sum_{k=1}^{L} \eta_k(\xi) f_k$ defines a tensor $F \in (\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ such that

$$F = \sum_{k=1}^{L} f_k \otimes \eta_k$$

Tensor structured equations

- $\lambda : \Xi \to \mathbb{R}$ can be identified with an operator $\Lambda : \mathcal{S} \to \widetilde{\mathcal{S}}'$ such that

$$\langle \Lambda s, \tilde{s} \rangle = \mathbb{E}_\mu(\lambda(\xi)s(\xi)\tilde{s}(\xi))$$

- $A(\xi) = \sum_{k=1}^R \lambda_k(\xi)A_k$ defines an operator B from $\mathcal{V} \otimes \mathcal{S}$ to $(\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ such that

$$B = \sum_{k=1}^R A_k \otimes \Lambda_k$$

- $f(\xi) = \sum_{k=1}^L \eta_k(\xi) f_k$ defines a tensor $F \in (\mathcal{W} \otimes \widetilde{\mathcal{S}})'$ such that

$$F = \sum_{k=1}^L f_k \otimes \eta_k$$

- Tensor structured equation

$$u \in \mathcal{V} \otimes \mathcal{S}, \quad Bu = F \quad \iff \quad \left(\sum_{k=1}^R A_k \otimes \Lambda_k \right) u = \sum_{k=1}^L f_k \otimes \eta_k$$

- For $\{\Phi_i\}_{i=1}^{\mathcal{P}}$ and $\{\Psi_i\}_{i=1}^{\mathcal{P}}$ bases of $\mathcal{S}$ and $\widetilde{\mathcal{S}}$, algebraic representation of Λ:

$$\boldsymbol{\Lambda} \in \mathbb{R}^{\mathcal{P}\times\mathcal{P}}, \quad (\boldsymbol{\Lambda})_{ij} = \langle \Lambda\Phi_j, \Psi_i\rangle = \mathbb{E}_\mu(\lambda(\xi)\Phi_j(\xi)\Psi_i(\xi))$$

- $u \in \mathcal{V}\otimes\mathcal{S}$ identified with a tensor $\mathbf{u} \in \mathbb{R}^{\mathcal{N}} \otimes \mathbb{R}^{\mathcal{P}}$ such that

$$u = \sum_{i=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{P}} (\mathbf{u})_{ij}\varphi_i \otimes \Phi_j$$

- Tensor structured equation in algebraic form

$$\mathbf{u} \in \mathbb{R}^{\mathcal{N}} \otimes \mathbb{R}^{\mathcal{P}}, \quad \mathbf{B}\mathbf{u} = \mathbf{F} \quad \Longleftrightarrow \quad \left(\sum_{k=1}^{R} \mathbf{A}_k \otimes \boldsymbol{\Lambda}_k\right)\mathbf{u} = \sum_{k=1}^{L} \mathbf{f}_k \otimes \boldsymbol{\eta}_k$$

Higher order tensor structure

- Suppose that $\mu = \mu_1 \otimes \dots \mu_d$, a product measure on $\Xi_1 \times \dots \times \Xi_d$ (e.g. when $\xi = (\xi_1, \dots, \xi_d)$ are independent random variables). Then

$$L^2_\mu(\Xi) = \overline{L^2_{\mu_1}(\Xi_1) \otimes \dots \otimes L^2_{\mu_d}(\Xi_d)}$$

- Suppose that the approximation space $\mathcal{S} \subset L^2_\mu(\Xi)$ is a finite dimensional tensor space

$$\mathcal{S} = \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_d, \quad \mathcal{S}_\nu \subset L^2_{\mu_\nu}(\Xi_\nu)$$

  and the same for $\widetilde{\mathcal{S}} = \widetilde{\mathcal{S}}_1 \otimes \dots \otimes \widetilde{\mathcal{S}}_d$.
- $\lambda^{(\nu)} : \Xi_\nu \to \mathbb{R}$ can be identified with an operator $\Lambda^{(\nu)} : \mathcal{S}_\nu \to \widetilde{\mathcal{S}}'_\nu$ such that

$$\langle \Lambda^{(\nu)} s, \tilde{s} \rangle = \mathbb{E}_{\mu_\nu}(\lambda^{(\nu)}(\xi_\nu) s(\xi_\nu) \tilde{s}(\xi_\nu))$$

- A function $\lambda : \Xi \to \mathbb{R}$ such that $\lambda(\xi) = \lambda^{(1)}(\xi_1) \dots \lambda^{(d)}(\xi_d)$ can be identified with an operator $\Lambda = \mathcal{S} \to \widetilde{\mathcal{S}}'$ such that

$$\Lambda = \Lambda^{(1)} \otimes \dots \otimes \Lambda^{(d)}$$

- Suppose that

$$A(\xi) = \sum_{k=1}^{R} A_k \lambda_k(\xi), \quad \text{with} \quad \lambda_k(\xi) = \lambda_k^{(1)}(\xi_1)\dots\lambda_k^{(d)}(\xi_d)$$

and

$$f(\xi) = \sum_{k=1}^{L} f_k \eta_k(\xi), \quad \text{with} \quad \eta_k(\xi) = \eta_k^{(1)}(\xi_1)\dots\eta_k^{(d)}(\xi_d)$$

- Tensor structured equation for $u \in \mathcal{V} \otimes \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_d$

$$Bu = F \quad \Longleftrightarrow \quad \left(\sum_{k=1}^{R} B_k \otimes \Lambda_k^{(1)} \otimes \dots \otimes \Lambda_k^{(d)}\right) u = \sum_{k=1}^{L} f_k \otimes \eta_k^{(1)} \otimes \dots \otimes \eta_k^{(d)}$$

- Tensor structured equation in algebraic form for $\mathbf{u} \in \mathbb{R}^{\mathcal{N}} \otimes \mathbb{R}^{\mathcal{P}_1} \otimes \dots \otimes \mathbb{R}^{\mathcal{P}_d}$

$$\mathbf{B}\mathbf{u} = \mathbf{F} \quad \Longleftrightarrow \quad \left(\sum_{k=1}^{R} \mathbf{A}_k \otimes \boldsymbol{\Lambda}_k^{(1)} \otimes \dots \otimes \boldsymbol{\Lambda}_k^{(d)}\right) \mathbf{u} = \sum_{k=1}^{L} \mathbf{f}_k \otimes \boldsymbol{\eta}_k^{(1)} \otimes \dots \otimes \boldsymbol{\eta}_k^{(d)}$$

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Order-two tensors

- For an order-two tensor $w \in \mathcal{V} \otimes \mathcal{S}$, single notion of rank:

$$rank(w) \leq m \quad \Leftrightarrow \quad w = \sum_{i=1}^{m} v_i \otimes s_i$$

Order-two tensors

- For an order-two tensor $w \in \mathcal{V} \otimes \mathcal{S}$, single notion of rank:

$$rank(w) \leq m \quad \Leftrightarrow \quad w = \sum_{i=1}^{m} v_i \otimes s_i$$

- Set of rank-m tensors

$$\mathcal{R}_m = \{w \in \mathcal{V} \otimes \mathcal{S} : rank(w) \leq m\}$$

Order-two tensors

- For an order-two tensor $w \in \mathcal{V} \otimes \mathcal{S}$, single notion of rank:

$$rank(w) \leq m \quad \Leftrightarrow \quad w = \sum_{i=1}^{m} v_i \otimes s_i$$

- Set of rank-m tensors

$$\mathcal{R}_m = \{w \in \mathcal{V} \otimes \mathcal{S} : rank(w) \leq m\}$$

- Assume $\mathcal{V}$ is a Hilbert space. Then $w \in \mathcal{R}_m$ is identified with a rank-m operator $w : \mathcal{V} \rightarrow \mathcal{S}$ defined by

$$w(v) = \sum_{i=1}^{m} s_i \langle v_i, v \rangle, \quad Im(w) \subset span\,\{s_i\}_{i=1}^{m}$$

  The algebraic tensor space $\mathcal{V} \otimes \mathcal{S}$ is identified with the space $\mathcal{F}(\mathcal{V}, \mathcal{S})$ of finite rank operators from $\mathcal{V}$ to $\mathcal{S}$.

Order-two tensors

- For an order-two tensor $w \in \mathcal{V} \otimes \mathcal{S}$, single notion of rank:

$$rank(w) \leq m \quad \Leftrightarrow \quad w = \sum_{i=1}^{m} v_i \otimes s_i$$

- Set of rank-m tensors

$$\mathcal{R}_m = \{w \in \mathcal{V} \otimes \mathcal{S} : rank(w) \leq m\}$$

- Assume $\mathcal{V}$ is a Hilbert space. Then $w \in \mathcal{R}_m$ is identified with a rank-m operator $w : \mathcal{V} \to \mathcal{S}$ defined by

$$w(v) = \sum_{i=1}^{m} s_i \langle v_i, v \rangle, \quad Im(w) \subset span\,\{s_i\}_{i=1}^{m}$$

  The algebraic tensor space $\mathcal{V} \otimes \mathcal{S}$ is identified with the space $\mathcal{F}(\mathcal{V}, \mathcal{S})$ of finite rank operators from $\mathcal{V}$ to $\mathcal{S}$.

- For $w \in \mathbb{R}^{\mathcal{N}} \otimes \mathbb{R}^{\mathcal{P}}$, $w \in \mathcal{R}_m$ identified with a rank-m matrix

$$w = \sum_{i=1}^{m} v_i \otimes s_i \equiv \sum_{i=1}^{m} v_i s_i^T \in \mathbb{R}^{\mathcal{N} \times \mathcal{P}}$$

Order-two tensors

- For an order-two tensor $w \in \mathcal{V} \otimes \mathcal{S}$, single notion of rank:

$$rank(w) \leq m \quad \Leftrightarrow \quad w = \sum_{i=1}^{m} v_i \otimes s_i$$

- Set of rank-m tensors

$$\mathcal{R}_m = \{w \in \mathcal{V} \otimes \mathcal{S} : rank(w) \leq m\}$$

- Assume $\mathcal{V}$ is a Hilbert space. Then $w \in \mathcal{R}_m$ is identified with a rank-m operator $w : \mathcal{V} \to \mathcal{S}$ defined by

$$w(v) = \sum_{i=1}^{m} s_i \langle v_i, v \rangle, \quad Im(w) \subset span\{s_i\}_{i=1}^{m}$$

  The algebraic tensor space $\mathcal{V} \otimes \mathcal{S}$ is identified with the space $\mathcal{F}(\mathcal{V}, \mathcal{S})$ of finite rank operators from $\mathcal{V}$ to $\mathcal{S}$.

- For $w \in \mathbb{R}^{\mathcal{N}} \otimes \mathbb{R}^{\mathcal{P}}$, $w \in \mathcal{R}_m$ identified with a rank-m matrix

$$w = \sum_{i=1}^{m} v_i \otimes s_i \equiv \sum_{i=1}^{m} v_i s_i^T \in \mathbb{R}^{\mathcal{N} \times \mathcal{P}}$$

- $\mathcal{V} \otimes_{\|\cdot\|_\vee} \mathcal{S}$ equipped with the operator norm $\|\cdot\|_\vee$ (injective tensor norm) is identified with the closure of $\mathcal{F}(\mathcal{V}, \mathcal{S})$, which is the set of compact operators from $\mathcal{V}$ to $\mathcal{S}$.

Low-rank approximation of order-two tensors

Consider a tensor

$$u \in \mathcal{V} \otimes_{\|\cdot\|} \mathcal{S}$$

- Rank-m approximation of u

$$u_m = \sum_{i=1}^{m} v_i \otimes s_i$$

- Best rank-m approximation with respect to $\|\cdot\|$:

$$\|u - u_m\| = \min_{w \in \mathcal{R}_m} \|u - w\|$$

- Best rank-m approximation with respect to a certain "distance" to solution:

$$\mathcal{E}(u, u_m) = \min_{w \in \mathcal{R}_m} \mathcal{E}(u, w)$$

- That defines optimal vectors $\{v_i\}_{i=1}^m$ and $\{s_i\}_{i=1}^m$ with respect to different criteria.

- Subspace-based parametrization of $\mathcal{R}_m$

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

or

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

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

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

or

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

- That defines sequences of optimal subspaces $\mathcal{V}_m$ and $\mathcal{S}_m$ (with respect to the chosen "distance"). For $u_m = \sum_{i=1}^m v_i \otimes s_i$, $\mathcal{V}_m = span\{v_i\}_{i=1}^m$ and $\mathcal{S}_m = span\{s_i\}_{i=1}^m$.

Hilbert setting: induced norm and SVD

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

$$\langle v \otimes s, v' \otimes s' \rangle = \langle v, v' \rangle_{\mathcal{V}} \langle s, s' \rangle_{\mathcal{S}}.$$

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

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

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

$$u_m = \sum_{i=1}^{m} \sigma_i v_i \otimes s_i$$

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

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

- For $\mathcal{V} = \mathbb{R}^{\mathcal{N}}$ and $\mathcal{S} = \mathbb{R}^{\mathcal{P}}$, the canonical norm $\|\cdot\|$ coincides with the matrix Frobenius norm and

$$u = \sum_{i=1}^{min(\mathcal{N},\mathcal{P})} \sigma_i v_i \otimes s_i = V \Sigma S^T$$

with $V \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ and $S \in \mathbb{R}^{\mathcal{P}\times\mathcal{P}}$ orthogonal matrices and $\Sigma \in \mathbb{R}^{\mathcal{N}\times\mathcal{P}}$ a diagonal matrix.

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

- Natural (induced) norm

$$\|u\|_2 = \left(\int_\Xi \|u(\xi)\|^2_{\mathcal{V}} \mu(d\xi) \right)^{1/2}$$

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

- Natural (induced) norm

$$\|u\|_2 = \left(\int_\Xi \|u(\xi)\|^2_{\mathcal{V}} \mu(d\xi) \right)^{1/2}$$

- A rank-m approximation u_m is of the form

$$u_m(\xi) = \sum_{i=1}^{m} v_i s_i(\xi)$$

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

- Natural (induced) norm

$$\|u\|_2 = \left(\int_\Xi \|u(\xi)\|^2_\mathcal{V} \mu(d\xi) \right)^{1/2}$$

- A rank-m approximation u_m is of the form

$$u_m(\xi) = \sum_{i=1}^m v_i s_i(\xi)$$

- The best rank-r approximation u_r which solves

$$\min_{v \in \mathcal{R}_m} \|u - v\|_2^2 = \min_{dim(\mathcal{V}_m)=m} \|u - P_{\mathcal{V}_m} u\|_2^2$$

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

$$U\lambda = \int_\Xi u(\xi)\lambda(\xi)\mu(d\xi)$$

also known as Karhunen-Loeve decomposition for μ a probability measure (or Proper Orthogonal Decomposition for ξ the time)

Low-rank 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 $\|\cdot\|_\mathcal{V}$.

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

- Optimal rank-m approximation is defined by

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

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

- Set of solutions $\mathcal{K} = \{u(\xi); \xi \in \Xi\} \subset \mathcal{V}$. Assuming $\mathcal{K}$ compact,

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

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

- In general, optimal spaces are such that

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

Low-rank 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 $\|\cdot\|_\mathcal{V}$.

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

- Optimal rank-m approximation is defined by

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

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

- Set of solutions $\mathcal{K} = \{u(\xi); \xi \in \Xi\} \subset \mathcal{V}$. Assuming $\mathcal{K}$ compact,

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

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

- In general, optimal spaces are such that

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

- 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 $\|\cdot\|_\mathcal{V}$.

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

- Optimal rank-m approximation is defined by

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

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

- Set of solutions $\mathcal{K} = \{u(\xi); \xi \in \Xi\} \subset \mathcal{V}$. Assuming $\mathcal{K}$ compact,

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

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

- In general, optimal spaces are such that

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

Low-rank 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 $\|\cdot\|_\mathcal{V}$.

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

- Optimal rank-m approximation is defined by

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

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

- Set of solutions $\mathcal{K} = \{u(\xi); \xi \in \Xi\} \subset \mathcal{V}$. Assuming $\mathcal{K}$ compact,

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

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

- In general, optimal spaces are such that

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

Optimal low-rank approximation in the general case

In the general case (provided well-posedness of minimization problems), best rank-m approximation and corresponding optimal spaces are still well defined by

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

Optimal low-rank approximation in the general case

In the general case (provided well-posedness of minimization problems), best rank-m approximation and corresponding optimal spaces are still well defined by

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

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

$$u_m = \sum_{i=1}^{m} v_i^m \otimes s_i^m$$

How to recover a notion of decomposition ?

Suboptimal constructions with nested subspaces and notion of decomposition, based on greedy constructions of the approximation or of subspaces.

- Reduced Basis method (greedy algorithms) and Empirical Interpolation Method (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 $L^\infty \otimes L^\infty$)

Proper Generalized Decomposition

- Greedy construction of the approximation (well-known version of PGD)
  Starting from $u_0 = 0$, construction of a sequence $\{u_m\}_{m\geq 1}$ by successive corrections in the "dictionary" of rank-one elements $\mathcal{R}_1$:

$$\mathcal{E}(u, u_{m-1} + v_m \otimes s_m) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w)$$

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

Proper Generalized Decomposition

- Greedy construction of the approximation (well-known version of PGD)
  Starting from $u_0 = 0$, construction of a sequence $\{u_m\}_{m\geq 1}$ by successive corrections in the "dictionary" of rank-one elements $\mathcal{R}_1$:

$$\mathcal{E}(u, u_{m-1} + v_m \otimes s_m) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w)$$

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

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

$$\mathcal{E}(u, u_m) = \min_{\substack{dim(\mathcal{V}_m)=m \\ \mathcal{V}_m \supset \mathcal{V}_{m-1}}} \min_{\substack{dim(\mathcal{S}_m)=m \\ \mathcal{S}_m \supset \mathcal{S}_{m-1}}} \min_{w \in \mathcal{V}_m \otimes \mathcal{S}_m} \mathcal{E}(u, w) = \min_{v_m \in \mathcal{V}} \min_{s_m \in \mathcal{S}} \min_{\sigma \in \mathbb{R}^{m \times m}} \mathcal{E}(u, \sum_{i,j=1}^{m} \sigma_{ij} v_i \otimes s_j)$$

  or partially greedy construction of subspaces

$$\mathcal{E}(u, u_m) = \min_{\substack{dim(\mathcal{V}_m)=m \\ \mathcal{V}_m \supset \mathcal{V}_{m-1}}} \min_{w \in \mathcal{V}_m \otimes \mathcal{S}} \mathcal{E}(u, w) = \min_{v_m \in \mathcal{V}} \min_{\{s_i\}_{i=1}^m} \mathcal{E}(u, \sum_{i=1}^{m} v_i \otimes s_i)$$

- 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 [N. 2007]

$$\mathcal{E}(u, u_{m-1} + v_m \otimes s_m) = \min_{v \in \mathcal{V}} \min_{s \in \mathcal{S}} \mathcal{E}(u, u_{m-1} + v \otimes s)$$

$$\mathcal{E}(u, u_m) = \min_{w \in \mathcal{V}_m \otimes \mathcal{S}} \mathcal{E}(u, w), \quad \text{with} \quad \mathcal{V}_m = span\{v_i\}_{i=1}^m$$

$$u_m = \sum_{i=1}^m v_i \otimes s_i^m$$

Greedy construction of a reduced basis $\{v_1, \dots, v_m, \dots\}$.

Remark : Convergence results are available but still no a priori estimates.

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Classical iterative methods with low-rank truncations

- Equation in tensor format

$$Bu = F$$

- Iterative solver (Richardson, Gradient...)

$$u^{(k)} = T(u^{(k-1)}) \quad (T\text{: iteration map})$$

For example

$$u^{(k)} = u^{(k-1)} - \alpha(Bu^{(k-1)} - F)$$

- Approximate iterations using low-rank truncations:

$$u^{(k)} \in \mathcal{R}_{m(\epsilon)} \quad \text{such that} \quad \|u^{(k)} - T(u^{(k-1)})\| \leq \epsilon$$

- For the canonical norm $\|\cdot\|$, truncation based on SVD
- Computational requirements: low-rank algebra and efficient SVD algorithms.
- Analysis : perturbation of iterative algorithms.

(see [Matthies and Zander 2012])

Minimal residual low-rank approximation

- Tensor structured equation
$$Bu = F$$
- Residual-based error
$$\mathcal{E}(u, w) = \|Bw - F\|_C = \|w - u\|_{B^*CB}$$
with a certain residual norm $\|\cdot\|_C^2 = \langle C\cdot, \cdot\rangle$.
- Best rank-m approximation
$$\mathcal{E}(u, u_m) = \min_{w \in \mathcal{R}_m} \mathcal{E}(u, w)$$

- Assuming $\tilde{\alpha}\|w\| \leq \|w\|_{B^*CB} \leq \tilde{\gamma}\|w\|$, then quasi-optimal approximation:

$$\|u - u_m\| \leq \frac{1}{\tilde{\alpha}}\|Bu_m - F\|_C = \frac{1}{\tilde{\alpha}} \min_{w \in \mathcal{R}_m} \|Bw - F\|_C \leq \frac{\tilde{\gamma}}{\tilde{\alpha}} \min_{w \in \mathcal{R}_m} \|u - w\|$$

- Importance of well-conditioned formulations, with $\frac{\tilde{\gamma}}{\tilde{\alpha}} \approx 1$.
- Construction of preconditioners in low-rank format [Giraldi 2012]
- Goal-oriented approach by choosing C such that

$$\|Bw - F\|_C = \|w - u\|_\star$$

where $\| \cdot \|_\star$ is a norm constructed by taking into account the objective of the computation [PhD O. Zahm]

- Assuming $\tilde{\alpha}\|w\| \leq \|w\|_{B^*CB} \leq \tilde{\gamma}\|w\|$, then quasi-optimal approximation:

$$\|u - u_m\| \leq \frac{1}{\tilde{\alpha}}\|Bu_m - F\|_C = \frac{1}{\tilde{\alpha}} \min_{w\in\mathcal{R}_m} \|Bw - F\|_C \leq \frac{\tilde{\gamma}}{\tilde{\alpha}} \min_{w\in\mathcal{R}_m} \|u - w\|$$

- Importance of well-conditioned formulations, with $\frac{\tilde{\gamma}}{\tilde{\alpha}} \approx 1$.
- Construction of preconditioners in low-rank format [Giraldi 2012]
- Goal-oriented approach by choosing C such that

$$\|Bw - F\|_C = \|w - u\|_\star$$

where $\|\cdot\|_\star$ is a norm constructed by taking into account the objective of the computation [PhD O. Zahm]

Remark: another residual-based error

$$\mathcal{E}(u,w)^2 = \mathbb{E}_\mu(\|A(\xi)w(\xi) - f(\xi)\|^2_{D(\xi)}) = \mathbb{E}_\mu(\|w(\xi) - u(\xi)\|^2_{A(\xi)^*D(\xi)A(\xi)})$$

with a certain residual norm $\|\cdot\|_{D(\xi)}$ on $\mathcal{W}'$. For symmetric problems and $D(\xi) = A(\xi)^{-1}$, it yields $\mathcal{E}(u,w) = \|Bw - F\|_{B^{-1}}$.

PGD algorithm in practice

- Ideal rank-m approximation u_m defined by

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

- Supoptimal greedy construction of subspaces $\mathcal{V}_m$: Starting from $\mathcal{V}_0 = 0$, we define a sequence of rank-m approximations u_m by

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

  Denoting $u_m = \sum_{i=1}^m v_i \otimes s_i^m$, we have

$$\mathcal{E}(u,\sum_{i=1}^m v_i\otimes s_i^m) = \min_{v_m\in\mathcal{V}} \min_{(s_1,\ldots,s_m)\in\mathcal{S}^m} \mathcal{E}(u,\sum_{i=1}^m v_i\otimes s_i) \quad (1)$$

- Alternating minimization algorithm for solving (1): solve successively

$$\min_{v_m \in \mathcal{V}} \mathcal{E}(u, \sum_{i=1}^{m} v_i \otimes s_i)^2, \tag{2}$$

$$\min_{(s_1,\dots,s_m)\in\mathcal{S}^m} \mathcal{E}(u, \sum_{i=1}^{m} v_i \otimes s_i)^2 \tag{3}$$

- Consider a symmetric problem, and let $C = B^{-1}$ so that

$$\mathcal{E}(u,w)^2 = \|Bw - F\|^2_{B^{-1}} = \langle Bw - F, w - u\rangle = \mathbb{E}_\mu\left(\langle A(\xi)w(\xi) - f(\xi), w(\xi) - u(\xi)\rangle\right)$$

- Solution of (2) (non parametric problem):

$$\min_{v_m \in \mathcal{V}} \|B \sum_{i=1}^{m} v_i \otimes s_i - F\|_{B^{-1}}^2 \quad \Leftrightarrow \quad \langle B \sum_{i=1}^{m} v_i \otimes s_i - F, \tilde{v} \otimes s_m \rangle = 0 \quad \forall \tilde{v} \in \mathcal{V}$$

which yields

$$\boxed{\widehat{A}_{mm} v_m = \widehat{f}_m - \sum_{i=1}^{m-1} \widehat{A}_{mi} v_i}$$

with

$$\widehat{A}_{mi} = \mathbb{E}_\mu(A(\xi) s_m(\xi) s_i(\xi)) = \sum_{k=1}^{R} A_k \widehat{\lambda}_{k,m,i}, \quad \widehat{\lambda}_{k,m,i} = \mathbb{E}_\mu(\lambda_k(\xi) s_m(\xi) s_i(\xi))$$

$$\widehat{f}_m = \mathbb{E}_\mu(f(\xi) s_m(\xi)) = \sum_{k=1}^{L} f_k \widehat{\eta}_{k,m}, \quad \widehat{\eta}_{k,m} = \mathbb{E}_\mu(\eta_k(\xi) s_m(\xi))$$

- $\widehat{A}_{mi}$ is an evaluation of $A(\xi) = \sum_{k=1}^{R} A_k \lambda_k(\xi)$ for particular values of the λ_k.
- $\widehat{f}_m$ is an evaluation of $f(\xi) = \sum_{k=1}^{L} f_k \eta_k(\xi)$ for particular values of the η_k.
- It looks like a sampling approach but it is not ! (no sampling of ξ)

Example 1

$$\langle A(\xi)v, w\rangle = \int_D \nabla w(x) \cdot \kappa(x,\xi) \cdot \nabla v(x) dx, \quad \langle f(\xi), w\rangle = \int_D g(x,\xi) w(x) dx$$

- $\langle \widehat{A}_{mi} v, w\rangle = \int_D \nabla w(x) \cdot \widehat{\kappa}_{mi} \cdot \nabla v(x) dx$ with $\widehat{\kappa}_{mi}(x) = \mathbb{E}_\mu(\kappa(x,\xi) s_m(\xi) s_i(\xi))$
- $\langle \widehat{f}_m, w\rangle = \int_D \widehat{g}_m(x) w(x) dx$ with $\widehat{g}_m(x) = \mathbb{E}_\mu(g(x,\xi) s_m(\xi))$

- Solution of (3) (reduced order parametric problem):

$$\min_{(s_1,\dots,s_m)\in\mathcal{S}^m} \|B\sum_{i=1}^m v_i \otimes s_i - F\|_{B^{-1}}^2$$

Denoting $\mathbf{s} = (s_i)_{i=1}^m \in (\mathcal{S})^m$, it yields

$$\boxed{\mathbb{E}_\mu(\mathbf{t}(\xi)^T \mathbf{A}_m(\xi)\mathbf{s}(\xi)) = \mathbb{E}_\mu(\mathbf{t}(\xi)^T \mathbf{f}_m(\xi)) \quad \forall \mathbf{t} \in (\mathcal{S})^m} \tag{4}$$

with reduced parametrized matrix and vector

$$(\mathbf{A}_m(\xi))_{ij} = \langle A(\xi)v_j, v_i\rangle, \quad (\mathbf{f}_m(\xi))_i = \langle f(\xi), v_i\rangle.$$

Solution $\mathbf{s}(\xi)$ of (4) is the stochastic Galerkin approximation of the solution of

$$\boxed{\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi)} \tag{5}$$

- Solution of (3) (reduced order parametric problem):

$$\min_{(s_1,\dots,s_m)\in\mathcal{S}^m} \|B\sum_{i=1}^{m} v_i\otimes s_i - F\|^2_{B^{-1}}$$

Denoting $\mathbf{s}=(s_i)_{i=1}^m \in (\mathcal{S})^m$, it yields

$$\boxed{\mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{A}_m(\xi)\mathbf{s}(\xi)) = \mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{f}_m(\xi)) \quad \forall \mathbf{t}\in(\mathcal{S})^m} \tag{4}$$

with reduced parametrized matrix and vector

$$(\mathbf{A}_m(\xi))_{ij} = \langle A(\xi)v_j, v_i\rangle, \quad (\mathbf{f}_m(\xi))_i = \langle f(\xi), v_i\rangle.$$

Solution $\mathbf{s}(\xi)$ of (4) is the stochastic Galerkin approximation of the solution of

$$\boxed{\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi)} \tag{5}$$

  - Using low-rank (affine) representations of $A(\xi)$ and $f(\xi)$, we obtain

$$\mathbf{A}_m(\xi) = \sum_{k=1}^{R}\mathbf{A}_{m,k}\lambda_k(\xi), \quad \mathbf{f}_m(\xi) = \sum_{k=1}^{L}\mathbf{f}_{m,k}\eta_k(\xi).$$

- Solution of (3) (reduced order parametric problem):

$$\min_{(s_1,\ldots,s_m)\in\mathcal{S}^m} \|B\sum_{i=1}^m v_i \otimes s_i - F\|^2_{B^{-1}}$$

Denoting $\mathbf{s} = (s_i)_{i=1}^m \in (\mathcal{S})^m$, it yields

$$\boxed{\mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{A}_m(\xi)\mathbf{s}(\xi)) = \mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{f}_m(\xi)) \quad \forall \mathbf{t} \in (\mathcal{S})^m} \tag{4}$$

with reduced parametrized matrix and vector

$$(\mathbf{A}_m(\xi))_{ij} = \langle A(\xi)v_j, v_i\rangle, \quad (\mathbf{f}_m(\xi))_i = \langle f(\xi), v_i\rangle.$$

Solution $\mathbf{s}(\xi)$ of (4) is the stochastic Galerkin approximation of the solution of

$$\boxed{\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi)} \tag{5}$$

  - Using low-rank (affine) representations of $A(\xi)$ and $f(\xi)$, we obtain

$$\mathbf{A}_m(\xi) = \sum_{k=1}^R \mathbf{A}_{m,k}\lambda_k(\xi), \quad \mathbf{f}_m(\xi) = \sum_{k=1}^L \mathbf{f}_{m,k}\eta_k(\xi).$$

  - (4) is a system of $m \times \dim(\mathcal{S})$ equations. If $\dim(\mathcal{S}) \gg 1$, structured approximation in $\mathcal{S}$ can be used to reduced the cost (sparsity, low-rank...).

- Solution of (3) (reduced order parametric problem):

$$\min_{(s_1,\ldots,s_m)\in\mathcal{S}^m} \|B\sum_{i=1}^{m} v_i \otimes s_i - F\|^2_{B^{-1}}$$

Denoting $\mathbf{s} = (s_i)_{i=1}^m \in (\mathcal{S})^m$, it yields

$$\boxed{\mathbb{E}_\mu(\mathbf{t}(\xi)^T \mathbf{A}_m(\xi)\mathbf{s}(\xi)) = \mathbb{E}_\mu(\mathbf{t}(\xi)^T \mathbf{f}_m(\xi)) \quad \forall \mathbf{t} \in (\mathcal{S})^m} \tag{4}$$

with reduced parametrized matrix and vector

$$(\mathbf{A}_m(\xi))_{ij} = \langle A(\xi)v_j, v_i\rangle, \quad (\mathbf{f}_m(\xi))_i = \langle f(\xi), v_i\rangle.$$

Solution $\mathbf{s}(\xi)$ of (4) is the stochastic Galerkin approximation of the solution of

$$\boxed{\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi)} \tag{5}$$

- Using low-rank (affine) representations of $A(\xi)$ and $f(\xi)$, we obtain

$$\mathbf{A}_m(\xi) = \sum_{k=1}^{R} \mathbf{A}_{m,k}\lambda_k(\xi), \quad \mathbf{f}_m(\xi) = \sum_{k=1}^{L} \mathbf{f}_{m,k}\eta_k(\xi).$$

- (4) is a system of $m \times \dim(\mathcal{S})$ equations. If $\dim(\mathcal{S}) \gg 1$, structured approximation in $\mathcal{S}$ can be used to reduced the cost (sparsity, low-rank...).
- (5) can be solved with sampling-based approaches (interpolation, regularized least-squares...)

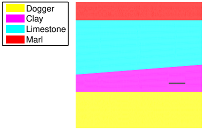
Example: stochastic Groundwater flow equation (MOMAS/Couplex)

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



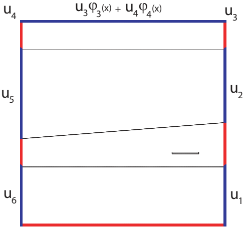
κ's probability laws

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



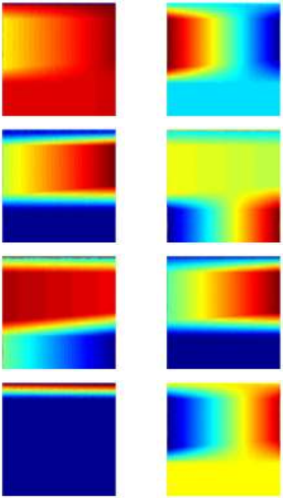
Neumann homogeneous

Dirichlet

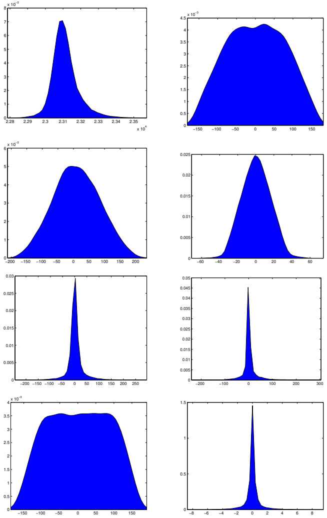
	Law
u_1	$U(288, 290)$
u_2	$U(305, 315)$
u_3	$U(330, 350)$
u_4	$U(170, 190)$
u_5	$U(195, 205)$
u_6	$U(285, 287)$

First modes with the greedy construction of the approximation

Spatial modes $\{v_1, \ldots, v_8\}$

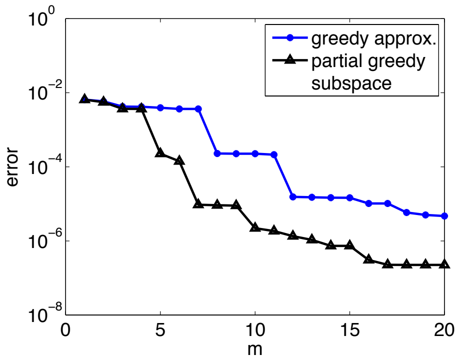


Stochastic modes $\{s_1, \ldots, s_8\}$: pdf



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

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



PGD based on Galerkin orthogonality criteria

- Approximation u_m in a subset $\mathcal{M}_m$
- For symmetric problems

$$\|Bu_m - F\|_{B^{-1}}^2 = \min_{w \in \mathcal{M}_m} \|Bw - F\|_{B^{-1}}^2 = \min_{w \in \mathcal{M}_m} \langle Bw - F, w - u \rangle$$

  Necessary (but not sufficient) condition of optimality

$$\langle Bu_m - F, \delta w \rangle = 0 \quad \forall \delta w \in T_{u_m}\mathcal{M}_m \tag{6}$$

  where $T_{u_m}\mathcal{M}_m$ is the tangent space to $\mathcal{M}_m$ at u_m.
- For more general problems (provided $B : \mathcal{V} \otimes \mathcal{S} \to (\mathcal{V} \otimes \mathcal{S})'$), search u_m in $\mathcal{M}_m$ such that it verifies (6).
- Heuristic approach. No theoretical results except for particular cases.

- For the greedy construction of the approximation

$$\mathcal{M}_m = u_{m-1} + \mathcal{R}_1, \quad u_m = u_{m-1} + v_m \otimes s_m,$$

$$T_{u_m}\mathcal{M}_m = T_{v_m \otimes s_m}\mathcal{R}_1 = \{\delta v \otimes s_m + v_m \otimes \delta s : \delta v \in \mathcal{V}, \delta s \in \mathcal{S}\}$$

$$(6) \Longleftrightarrow \langle B(u_{m-1} + v_m \otimes s_m) - F, \delta v \otimes s_m + v_m \otimes \delta s \rangle = 0 \quad \forall \delta v \in \mathcal{V}, \forall \delta s \in \mathcal{S}$$

- For the greedy construction of the approximation

$$\mathcal{M}_m = u_{m-1} + \mathcal{R}_1, \quad u_m = u_{m-1} + v_m \otimes s_m,$$

$$T_{u_m}\mathcal{M}_m = T_{v_m \otimes s_m}\mathcal{R}_1 = \{\delta v \otimes s_m + v_m \otimes \delta s : \delta v \in \mathcal{V}, \delta s \in \mathcal{S}\}$$

$$(6) \iff \langle B(u_{m-1} + v_m \otimes s_m) - F, \delta v \otimes s_m + v_m \otimes \delta s \rangle = 0 \quad \forall \delta v \in \mathcal{V}, \forall \delta s \in \mathcal{S}$$

- For the partial greedy construction of subspaces

$$\mathcal{M}_m = \{w \in \mathcal{V}_m \otimes \mathcal{S} : \dim(\mathcal{V}_m) = m, \mathcal{V}_m \supset \mathcal{V}_{m-1}\}$$
$$= \left\{ w = \sum_{i=1}^{m} v_i \otimes s_i : v_m \in \mathcal{V}, \{s_i\}_{i=1}^m \in \mathcal{S}^m \right\}$$

$$T_{u_m}\mathcal{M}_m = \left\{ \delta v \otimes s_m + \sum_{i=1}^{m} v_i \otimes \delta s_i : \delta v_m \in \mathcal{V}, \{\delta s_i\}_{i=1}^m \in \mathcal{S}^m \right\}$$

$$(6) \iff \langle Bu_m - F, \delta v \otimes s_m + \sum_{i=1}^{m} v_i \otimes \delta s_i \rangle = 0 \quad \forall \delta v \in \mathcal{V}, \ \forall \{\delta s_i\}_{i=1}^m \in \mathcal{S}^m$$

- Use of alternating direction algorithms. Computational aspects are similar to the case of symmetric problems described before.
  - non-parametric problems of the form

$$\widehat{A}_{mm} v_m = \widehat{f}_m - \sum_{i=1}^{m-1} \widehat{A}_{mi} v_i$$

  with

$$\widehat{A}_{mi} = \mathbb{E}_\mu(A(\xi) s_m(\xi) s_i(\xi)) = \sum_{k=1}^{R} A_k \widehat{\lambda}_{k,m,i}, \quad \widehat{\lambda}_{k,m,i} = \mathbb{E}_\mu(\lambda_k(\xi) s_m(\xi) s_i(\xi))$$

$$\widehat{f}_m = \mathbb{E}_\mu(f(\xi) s_m(\xi)) = \sum_{k=1}^{L} f_k \widehat{\eta}_{k,m}, \quad \widehat{\eta}_{k,m} = \mathbb{E}_\mu(\eta_k(\xi) s_m(\xi))$$

  - reduced order parametric problems of the form

$$\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi) \quad \text{or} \quad a_m(\xi) s_m(\xi) = g_m(\xi) - \sum_{i=1}^{m-1} a_{mi}(\xi) s_i(\xi)$$

- Possible application (heuristic approach) for nonlinear problems. B is a nonlinear map.

Illustration on a benchmark OPUS (http://www.opus-project.fr)

Cooling of electronic components

$$-\nabla(\kappa \cdot \nabla u) + D V \cdot \nabla u = f$$

- $\xi_1 = \kappa_{IC} \sim \log\mathcal{U}\big(0.2, 2\big) \rightsquigarrow$ diffusion coefficient
- $\xi_2 = r \sim \log\mathcal{U}\big(0.1, 100\big) \rightsquigarrow$ thermal contact conductance
- $\xi_3 = D \sim \log\mathcal{U}\big(5.10^{-4}, 10^{-2}\big) \rightsquigarrow$ advection field

Quantity of interest :

$$S_{\kappa_{IC}} = \frac{\mathbb{V}\left(\mathbb{E}\left(\int_{\Omega_{IC}} u | \kappa_{IC}\right)\right)}{\mathbb{V}\left(\int_{\Omega_{IC}} u\right)}$$

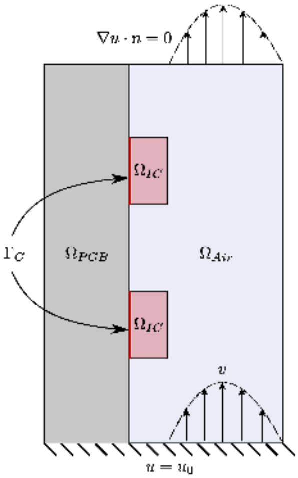


Figure : Geometry

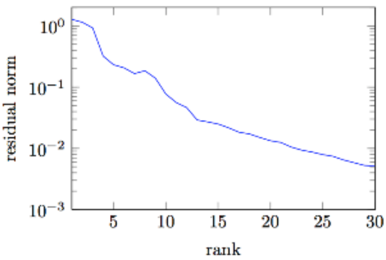


Figure : Evolution of the residual norm

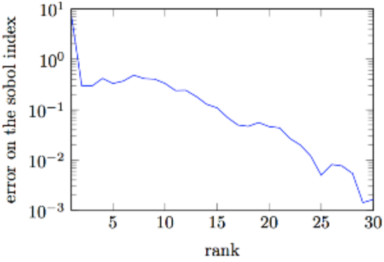
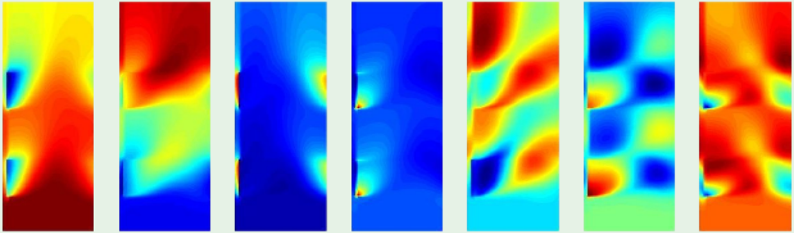


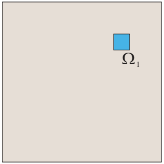
Figure : Error on the Sobol index

First spatial functions v_1 to v_7



Application to an advection-diffusion-reaction equation

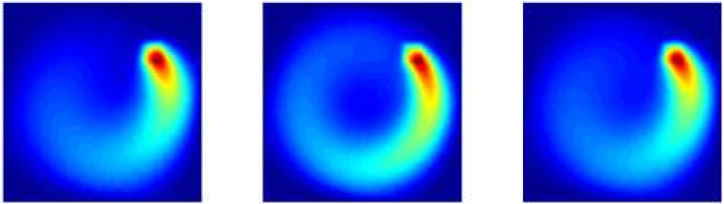
- $\partial_t u - a_1 \Delta u + a_2 c \cdot \nabla u + a_3 u = a_4 I_{\Omega_1}$ on $\Omega \times (0, T)$
- $u = 0$ on $\Omega \times \{0\}$
- $u = 0$ on $\partial\Omega \times (0, T)$



Uncertain parameters

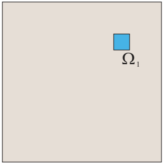
$$a_i(\xi) = \mu_{a_i}(1 + 0.2\xi_i), \quad \xi_i \in U(-1, 1), \quad \boxed{\Xi = (-1, 1)^4}$$

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



Application to an advection-diffusion-reaction equation

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



Uncertain parameters

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

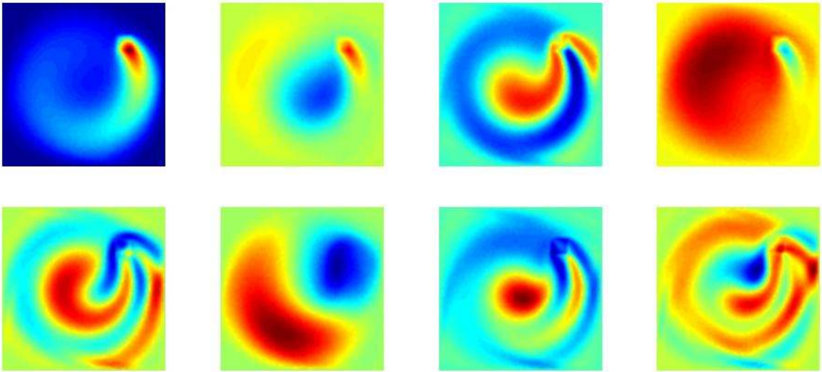
Low-rank approximation of the solution

$$\boxed{u(x, t, \xi) \approx \sum_{i=1}^{m} v_i(x, t) s_i(\xi)}$$

- Galerkin framework using Time-Discontinuous Galerkin approximation.
- Partial greedy construction of subspaces : Arnoldi-type algorithm [N. 2008]

Partial greedy construction of subspaces $\mathcal{V}_m$ with Arnoldi-type construction

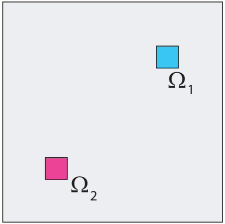
8 first modes of the decomposition $\{v_1(x,t) \ldots v_8(x,t)\}$



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

Quantity of interest

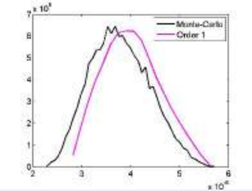
$$s(\xi) = \int_0^T \int_{\Omega_2} u(x,t,\xi)\, dxdt$$



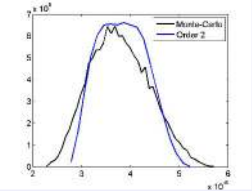
$$s_m(\xi) = \int_0^T \int_{\Omega_2} u_m(x,t,\xi)\, dxdt$$

Probability density function of $s_m(\xi)$

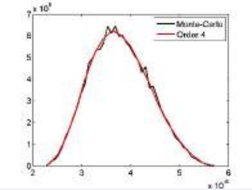
$m = 1$



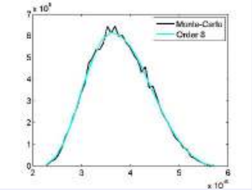
$m = 2$



$m = 4$



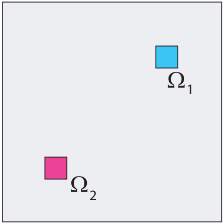
$m = 8$



Convergence of quantities of interest
Quantiles

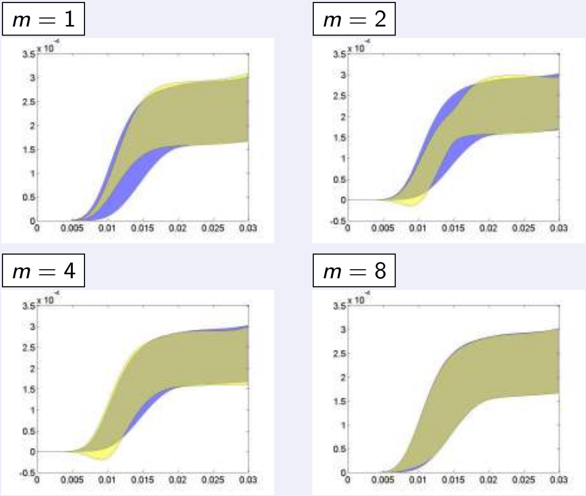
Quantity of interest

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



$$s_m(t,\xi) = \int_{\Omega_2} u_m(x,t,\xi)\,dx$$

99% Quantiles of $s_m(t,\xi)$



Low-rank approximation using sampling-based approach

- We want to compute an approximation of the solution $u(\xi)$, and then a variable of interest $s(u(\xi);\xi)$, for a collection of samples

$$\{\xi^k\}_{k=1}^K = \Xi_K$$

- The computation of

$$u(\xi^k) = B(\xi^k)^{-1} f(\xi^k) \quad \text{for all} \quad k = 1, \dots, K$$

  is unaffordable.

- Use of low-rank approximations ?

Low-rank approximation using sampling-based approach

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

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

- The best rank-m approximation u_m which solves

$$\min_{w \in \mathcal{R}_m} \|u - w\|_{2,K}^2 = \min_{w \in \mathcal{R}_m} \frac{1}{K} \sum_{k=1}^{K} \|u(\xi^k) - w(\xi^k)\|_{\mathcal{V}}^2$$

  corresponds to the truncated singular value decomposition of the tensor

$$\mathbf{u} = \{u(\xi^k)\}_{k=1}^K \in \mathcal{V}^K = \mathcal{V} \otimes \mathbb{R}^K$$

  also known as Empirical Karhunen-Loeve decomposition.

- Requires the solution of K independent problems (Black box simulations)

$$u(\xi^k) = B(\xi^k)^{-1} f(\xi^k), \quad k = 1, \dots, K$$

Low-rank approximation using sampling-based approach

- Residual based error

$$\mathcal{E}(u,w)^2 = \frac{1}{K}\sum_{k=1}^{K} \|A(\xi^k)w(\xi^k) - f(\xi^k)\|^2_{D(\xi^k)} = \|w-u\|^2_{\tilde{A},2,K}$$

- Best rank-m approximation defined by

$$\mathcal{E}(u,u_m) = \min_{w\in\mathcal{R}_m} \mathcal{E}(u,w)$$

- The restriction of $u:\Xi\to\mathcal{V}$ to Ξ_K is identified with the tensor

$$\mathbf{u} = \{u(\xi^k)\}_{k=1}^K \in (\mathcal{V})^K = \mathcal{V}\otimes\mathbb{R}^K$$

- $\|\cdot\|^2_{\tilde{A},2,K}$ defines on $\mathcal{V}\otimes\mathbb{R}^K$ a norm which is equivalent to $\|\cdot\|_{2,K}$ and

$$\|u-u_m\|_{2,K} \leq \frac{\tilde{\gamma}}{\tilde{\alpha}} \min_{v\in\mathcal{R}_m} \|u-v\|_{2,K}$$

- Denoting $\widehat{\mathbb{E}}^K_\mu(f(\xi)) = \frac{1}{K}\sum_{k=1}^K f(\xi^k)$,

$$\mathcal{E}(u,w)^2 = \widehat{\mathbb{E}}^K_\mu\left(\|A(\xi)w(\xi)-f(\xi)\|^2_{D(\xi)}\right)$$

- Low-rank algorithms follow the same lines simply replacing $\mathbb{E}_\mu(\cdot)$ by $\mathbb{E}^K_\mu(\cdot)$.

Outline

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Low-rank approximation for higher-order tensors

Approximation of a high-order tensor

$$u \in \overline{V_1 \otimes \ldots \otimes V_d}$$

in a subset of low-rank tensors

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

Rank of higher-order tensors ?

Higher-order low-rank formats

- Canonical rank:

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

Parametrization with $\sum_{\nu=1}^{d} \dim(V_\nu) = O(d)$ parameters.

Example

Ishigami function

$$v(x_1, x_2, x_3) = \sin(x_1) + a\sin(x_2)^2 + bx_3^4 \sin(x_1) = \sin(x_1)(1 + bx_3^4) + a\sin(x_2)^2$$

has a canonical rank 2.

Sparse tensor approximation

Suppose $\{\phi_i^\nu(x_\nu) : i \in \Lambda_\nu\}$ is a basis of V_ν (e.g. polynomial spaces). Then

$$\{\phi_{i_1}^1(x_1)\dots\phi_{i_d}^d(x_d) : (i_1,\dots,i_d) \in \Lambda = \Lambda_1 \times \dots \times \Lambda_d\}$$

is a basis of $V_1 \otimes \dots \otimes V_d$. An element $v \in\in V_1 \otimes \dots \otimes V_d$ can be written

$$v(x_1,\dots,x_d) = \sum_{i\in\Lambda} a_i \phi_{i_1}^1(x_1)\dots\phi_{i_d}^d(x_d)$$

A sparse tensor approximation v_N of v is under the form

$$v_N(x_1,\dots,x_d) = \sum_{i\in\Lambda_N} a_i \phi_{i_1}^1(x_1)\dots\phi_{i_d}^d(x_d)$$

with

$$\Lambda_N \subset \Lambda, \quad \#\Lambda_N = N \ll \#\Lambda$$

and has a canonical rank

$$rank(v_N) \leq N$$

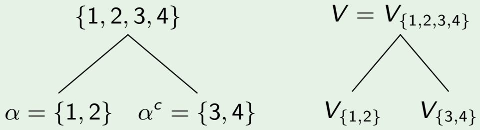
Consequence: if the best N-term approximation v_N (for a fixed basis) is such that $\|v - v_N\| = \sigma_N$, then

$$\inf_{v_N \in \mathcal{R}_N} \|v - v_N\| \leq \sigma_N$$

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

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

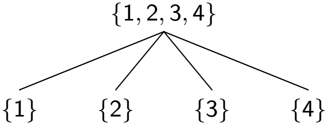
Example



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

- Tucker rank:

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



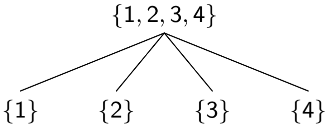
$$rank_T(v) \leq r = (r_1, \ldots, r_d) \iff v = \sum_{i_1=1}^{r_1} \ldots \sum_{i_d=1}^{r_d} a_{i_1 \ldots i_d} v^1_{i_1} \otimes \ldots \otimes v^d_{i_d}$$

- Tree-based Tucker rank:

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

- Tucker rank:

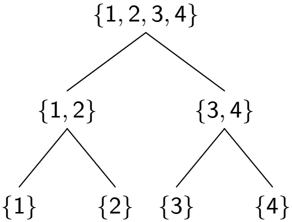
$$rank_T(v) = (rank_1(v), \dots, rank_d(v))$$



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

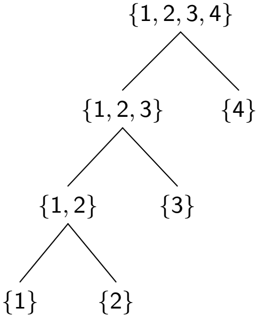
- Tree-based Tucker rank:

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



- TT-rank:

$$rank_{TT}(v) = (rank_{\{1\}}(v), rank_{\{1,2\}}(v) \ldots, rank_{\{1,\ldots,d-1\}}(v))$$



$$rank_{TT}(v) \leq r = (r_1, \ldots, r_{d-1}) \quad \Longleftrightarrow \quad v = \sum_{i_1=1}^{r_1} \ldots \sum_{i_d=1}^{r_{d-1}} v^1_{1,i_1} \otimes v^2_{i_1,i_2} \otimes \ldots \otimes v^d_{i_{d-1},1}$$

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 $v(x_1, \dots, x_d) = \sum_{i=1}^r \phi_i^1(x_1) \dots \phi_i^d(x_d)$ 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

1. Model reduction methods for high dimensional problems
2. Tensors and tensor-structured problems
3. Tensor-structured parametric and stochastic equations
4. Low-rank approximation of order-two tensors
5. Low-rank methods for parametric and stochastic equations
6. Low-rank approximation of higher order tensors
7. Higher-order low-rank methods for high-dimensional parametric and stochastic equations

Sample-based low-rank tensor approximation

- Computation of an output quantity of interest

$$s(\xi) = \ell(u(\xi); \xi), \quad \xi \sim \mu,$$

$$s \in L^2_\mu(\mathbb{R}^d)$$

Sample-based low-rank tensor approximation

- Computation of an output quantity of interest

$$s(\xi) = \ell(u(\xi); \xi), \quad \xi \sim \mu,$$

$$s \in L^2_\mu(\mathbb{R}^d)$$

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

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

Sample-based low-rank tensor approximation

- Computation of an output quantity of interest

$$s(\xi) = \ell(u(\xi); \xi), \quad \xi \sim \mu,$$

$$s \in L^2_\mu(\mathbb{R}^d)$$

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

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

- Suppose we have evaluations $\{s(\xi^k)\}_{k=1}^K$ of s at sample points $\{\xi^k\}_{k=1}^K$ (Simulations of the full order model or of a reduced order model).

Sample-based low-rank tensor approximation

- Computation of an output quantity of interest

$$s(\xi) = \ell(u(\xi); \xi), \quad \xi \sim \mu,$$

$$s \in L^2_\mu(\mathbb{R}^d)$$

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

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

- Suppose we have evaluations $\{s(\xi^k)\}_{k=1}^K$ of s at sample points $\{\xi^k\}_{k=1}^K$ (Simulations of the full order model or of a reduced order model).
- Low rank approximation using least-square minimization:

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

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

Sample-based low-rank tensor approximation

- Regularization could be required

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

Sample-based low-rank tensor approximation

- Regularization could be required

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

- For a given tensor format

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

solve

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

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

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

Sample-based low-rank tensor approximation

- Regularization could be required

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

- For a given tensor format

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

solve

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

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

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

Sample-based low-rank tensor approximation

- Regularization could be required

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

- For a given tensor format

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

solve

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

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

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

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

Now, entering the model...

Classical iterative methods with low-rank truncations

- Equation in tensor format

$$Bu = F, \quad u \in \mathcal{V} \otimes \mathcal{S}_1 \otimes \ldots \otimes \mathcal{S}_d$$

- Iterative solver

$$u^{(k)} = T(u^{(k-1)}) \quad (T\text{: iteration map})$$

- Approximate iterations using low-rank truncations:

$$u^{(k)} \in \mathcal{M}_{r(\epsilon)} \quad \text{such that} \quad \|u^{(k)} - T(u^{(k-1)})\| \leq \epsilon$$

- For the canonical norm $\|\cdot\|$, truncation based on higher-order SVD
- Computational requirements: low-rank algebra and efficient truncation algorithms
- Analysis : perturbation of iterative algorithms.

(see [Khoromskij & Schwab 2011])

Minimal residual low-rank approximation

- Residual-based error

$$\mathcal{E}(u, w) = \|Bw - F\|_C = \|w - u\|_{B^*CB}$$

with a certain residual norm $\| \cdot \|_C^2 = \langle C \cdot, \cdot \rangle$.

- Best approximation in $\mathcal{M}_{\leq r}$

$$\mathcal{E}(u, u_m) = \min_{w \in \mathcal{M}_{\leq r}} \mathcal{E}(u, w)$$

Approximation of higher-order tensors in canonical format

- Optimization in canonical format is ill-posed

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

- A quasi-optimal approximation $u_r \in \mathcal{R}_r$ can be obtained (but usually numerically unstable)

$$\mathcal{E}(u, u_r) \leq (1 + \epsilon) \inf_{v \in \mathcal{R}_r} \mathcal{E}(u, v)$$

- No notion of decomposition

$$u_r = \sum_{i=1}^{r} u_i^{(1,r)} \otimes \ldots \otimes u_i^{(d,r)}$$

Proper Generalized Decomposition in canonical format: greedy approximation

- Suboptimal construction of canonical representation using greedy algorithms.
- Starting from $u_0 = 0$, then
$$u_m = u_{m-1} + w_m$$
with $w_m = w_m^1 \otimes \ldots \otimes w_m^d$ defined by
$$\mathcal{E}(u, u_{m-1} + w_m) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w)$$
- Notion of decomposition
$$u_m = \sum_{i=1}^{m} w_i^1 \otimes \ldots \otimes w_i^d$$
- Possible optimization of functions after each correction for improving convergence (but we loose the notion of decomposition).

Proper Generalized Decomposition : greedy approximation

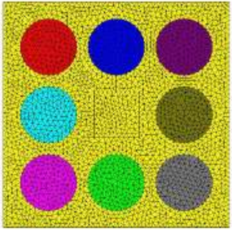
- Possible corrections w_m in other low-rank subsets.
- Convergence results available [Cances & al 2011, Falco & N. 2012] (not really specific to tensor setting) but no a priori estimates.
- Convergence may be slow compared to $\inf_{w \in \mathcal{R}_r} \mathcal{E}(u, w)$
- The construction does not really exploit the tensor structure.

A simple illustration on a diffusion equation

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

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

with $\xi_i \in U(-1,1)$



Approximation spaces

$$u \in \mathcal{V} \otimes \mathcal{S}, \quad \mathcal{S} = \mathbb{P}_{10}(-1,1) \otimes \ldots \otimes \mathbb{P}_{10}(-1,1)$$

Underlying approximation space with dimension 5.10^{11}

Residual-based error

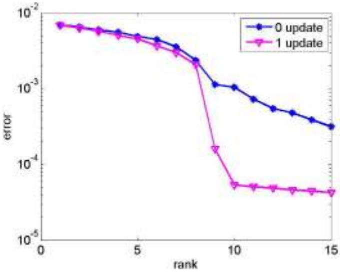
$$\mathcal{E}(u,v)^2 = \|v - u\|_B^2 = \mathcal{J}(v) - \mathcal{J}(u), \quad \mathcal{J}(v) = \int_\Omega \kappa \nabla v \cdot \nabla v - 2\int_\Omega I_D v$$

Progressive Galerkin PGD

$$u \approx u_m = \sum_{k=1}^{m} w_k, \quad w_k = v_k \otimes \phi_k^1 \otimes \ldots \otimes \phi_k^s \in \mathcal{R}_1$$

$$\min_{w_m \in \mathcal{R}_1} \|u - u_{m-1} - w_m\|_B^2 \quad + \quad \text{updates of functions } \{\phi_k^1, \ldots, \phi_k^s\}_{k=1}^m$$

Convergence (in L^2 norm)



Spatial modes w_k

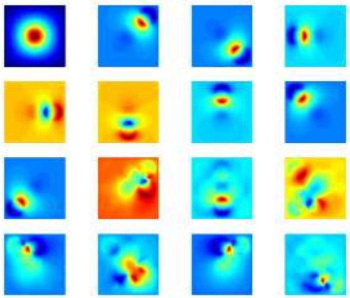
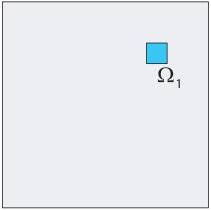


Illustration : stationary advection-diffusion-reaction equation

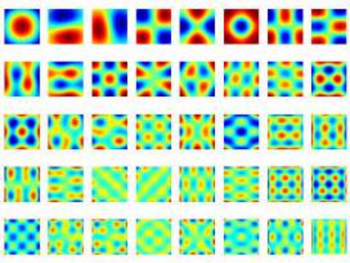
$$-\nabla \cdot (\kappa \nabla u) + c \cdot \nabla u + \gamma u = \delta I_{\Omega_1}(x) \quad on \quad \Omega$$



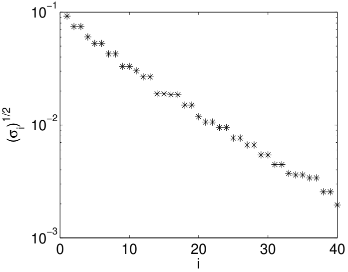
Random field

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

Spatial modes $\kappa_i(x)$



Amplitudes σ_i



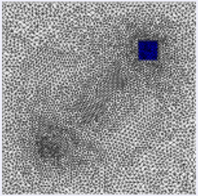
Stochastic approximation

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

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

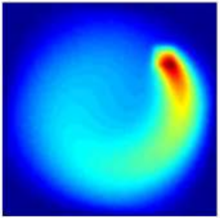
$$\boxed{dim(\mathcal{S}) = 5^{40} \approx 10^{28}}$$

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

$$u(\boldsymbol{\xi}) \approx u_m(\boldsymbol{\xi}) = \sum_{i=1}^{m} v_i s_i(\boldsymbol{\xi})$$

$$\hookrightarrow \quad \boxed{\mathcal{V}_m = span\{v_i\}_{i=1}^m}$$

low-rnak approximation of parametric functions

$$\mathbf{s}(\boldsymbol{\xi}) := (\mathbf{s}_i)_{i=1}^m \approx \mathbf{s}_Z(\boldsymbol{\xi}) = \sum_{k=1}^{Z} \phi_k^0 \prod_{j=1}^{d} \phi_k^j(\xi_j)$$

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

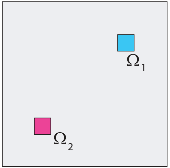
For a precision $\|u - u_{M,Z}\|_{L^2} \leqslant 10^{-2}$

- $\boxed{dim(\mathcal{V}_m) \approx 15} \ll 4435 = dim(\mathcal{V})$
- $\boxed{dim(\mathcal{S}_Z) \approx 10} \ll 10^{28} = dim(\mathcal{S})$
- 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

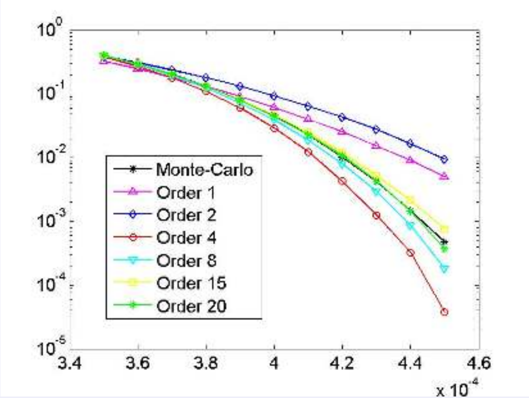
Quantity of interest

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



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

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



Convergence properties of quantities of interest

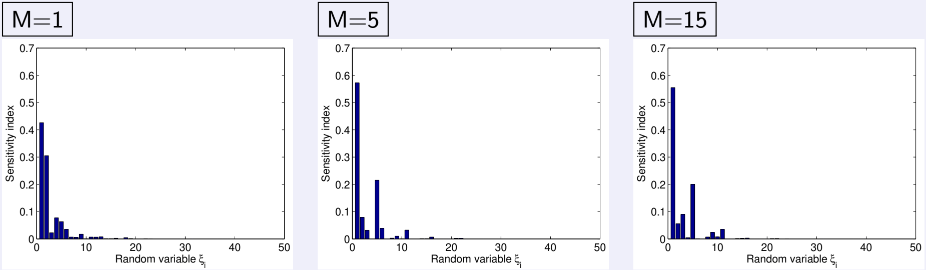
Sensitivity analysis

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

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



Proper Generalized Decomposition for higher-order tensors: Tucker format

- 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:

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

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

- 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

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

- 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

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

- 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

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

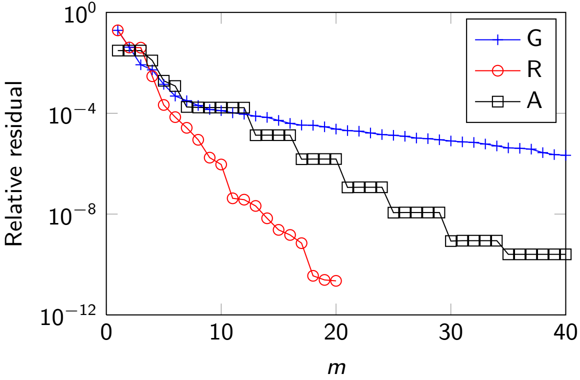
Illustration: PDE with random coefficients
with Loic Giraldi

$$
\begin{aligned}
& -\nabla \cdot (\nabla u) + \xi_2 u = 1 \quad on \quad D \backslash \omega \\
& -\nabla \cdot ((1+\xi_1)\nabla u) + \xi_2 u = 1 \quad on \quad \omega \\
& u = 0 \quad on \quad \partial\Omega
\end{aligned}
$$

$$
\begin{gathered}
\xi_1 \sim U(0,10), \quad \xi_2 \sim U(0,1) \\
u \in H^1_0(\Omega) \otimes L^2(\Xi_1) \otimes L^2(\Xi_2)
\end{gathered}
$$

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)

$\circ$ Suboptimal greedy construction of subspaces based on rank-one corrections (isotropic)

$\square$ Suboptimal greedy construction of subspaces with anisotropic enrichment

Illustration: PDE with random coefficients

Ranks with respect to iteration m for anisotropic construction

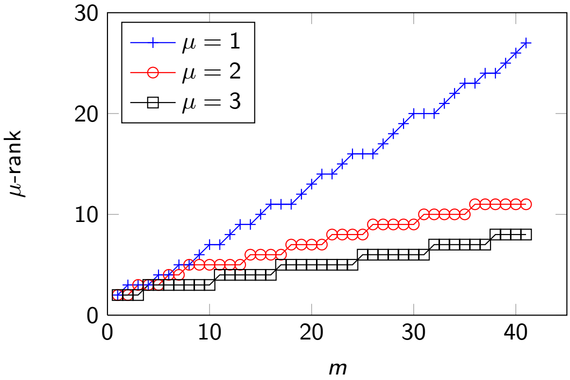
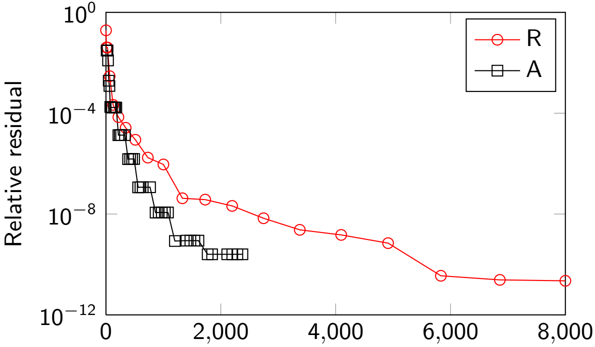


Illustration: PDE with random coefficients

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



$dim(U_m) = \prod_{\mu=1}^{3} dim(U_\mu^m)$

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

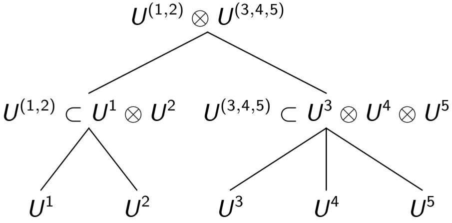
□ Suboptimal greedy construction of subspaces with anisotropic enrichment

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

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

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: variable of interest $s = \ell(u)$, 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.

B. Khoromskij. Tensors-structured numerical methods in scientific computing: Survey on recent advances. *Chemometrics and Intelligent Laboratory Systems*, 110(1):1 – 19, 2012.

A. Falco, W. Hackbusch and A. Nouy. Geometric structures 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.

References II

- **LOW-RANK FOR UQ AND PARAMETRIC MODELS**

A. Nouy. Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems. Archives Comp. Meth. Eng., 2010.

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

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.

H. G. Matthies and E. Zander. Solving stochastic systems with low-rank tensor compression. *Linear Algebra and its Applications*, 436(10), 2012.