Low-rank tensor methods for parametric and stochastic problems

Part 1: Low-rank methods and projection-based model reduction

Anthony Nouy

Ecole Centrale Nantes / GeM
Nantes, France

Stochastic and parametric analyses

Stochastic or parametric model

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

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

- Forward problem: given μ, compute a variable of interest

$$s(\xi) = g(u(\xi); \xi)$$

  and quantities of interest (statistical moments, probability of events, sensitivity indices...).
- Inverse problem: given observations of $s(\xi)$, determine ξ or estimate μ.
- Optimization: minimize objective function $s(\xi)$ over ξ.

Stochastic and parametric analyses

Ideal approach

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

Complexity issues

- Complex numerical models (**Part 1**)

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

$$\dim(\mathcal{V}) \gg 1$$

  - Limit the number of point evaluations
  - Remedy: projection-based model reduction, approximation of $u(\xi)$ in a low-dimensional subspace (or manifold) of $\mathcal{V}$

- Approximation of multivariate functions (**Part 2**)

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

$$d \gg 1 \text{ (possibly } d = \infty)$$

  - Classical approaches suffer from the curse of dimensionality
  - Remedy: adapted bases, structured approximations

A model example

Diffusion equations with random diffusion coefficient $\kappa(x, \omega)$:

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

- Groundwater flow (Nuclear Waste Disposal Simulation : Couplex)

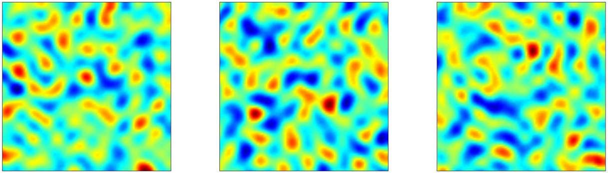
$$\kappa(x, \omega) = \sum_{i=1}^{d} \xi_i(\omega) I_{D_i}(x)$$

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	$\xi_4 \sim LU(10^{-5}, 10^{-4})$

3D problem requiring fine discretization : $\dim(\mathcal{V}) \gg 1$

- Random media with spatially correlated random fields

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



Outline

1. Functional framework for parametric and stochastic equations
2. Tensors
3. Low-rank approximation of order-two tensors
4. Computing low-rank approximations
5. Low-rank methods for parametric and stochastic equations

Outline

1. Functional framework for parametric and stochastic equations
2. Tensors
3. Low-rank approximation of order-two tensors
4. Computing low-rank approximations
5. Low-rank methods for parametric and stochastic equations

Notations, definitions

- ξ : parameters or vector-valued random variable with probability law μ.
- $\Xi \subset \mathbb{R}^d$: range of ξ (parameter set)
- μ : finite measure on Ξ
- 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

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

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

- Lebesgue space $L^p_\mu(\Xi) = L^p_\mu(\Xi; \mathbb{R})$
- $\mathbb{E}_\mu(v(\xi)) = \int_\Xi v(y)\mu(dy)$ (expectation)
- For X a normed vector space, X' denotes the algebraic dual of X and X^* denotes the topological dual of X.

Abstract formulation of a class of linear problems

Parametric (or stochastic) strong form

Find $u(\xi) \in \mathcal{V}$ such that it holds μ-almost surely

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

with $a(\cdot, \cdot; \xi) : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$ a bilinear form and $f(\cdot; \xi) : \mathcal{W} \to \mathbb{R}$ a continuous linear form.

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

- Uniformly continuous

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

- Uniformly weakly coercive

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

-

$$\forall w \in \mathcal{W} \setminus \{0\}, \quad \sup_{v \in \mathcal{V}} a(v, w) > 0$$

Examples

Example 1: diffusion equation with random diffusion coefficient

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

- $a(u,w;\xi) = \int_D \nabla w(x) \cdot \kappa(x,\xi) \cdot \nabla u(x)\, dx, \quad f(w;\xi) = \int_D g(x,\xi) w(x)\, dx$
- Approximation space $\mathcal{V} \subset H_0^1(D)$, $\mathcal{W} = \mathcal{V}$.
- $\alpha_\star \leq \kappa(x,\xi) \leq \gamma_\star$ for almost all x and ξ.
- $g(\cdot,\xi) \in L^2(\Omega)$.

Examples

Example 2: evolution equation

$$\partial_t u - \nabla \cdot (\kappa \nabla u) = g \quad \text{on } D \times I$$
$$u = u_0(\cdot, \xi) \text{ on } D \times \{0\}, \quad u = 0 \quad \text{on } \partial D \times I$$

- $\mathcal{V} \subset L^2(I; H^1_0(D)) \cap H^1(I; L^2(D))$ equipped with norm $\|v\|^2_{\mathcal{V}} = \|v\|^2_{L^2(I;H^1_0(D))} + \|v\|^2_{H^1(I;L^2(D))}$.
- $\mathcal{W} = \mathcal{W}_1 \times \mathcal{W}_2 \subset L^2(I; H^1_0(D)) \times L^2(D)$ equipped with norm $\|w\|^2_{\mathcal{W}} = \|w_1\|^2_{L^2(I;H^1_0(D))} + \|w_2\|^2_{L^2(D)}$.
- Bilinear and linear forms

$$a(v, w; \xi) = \int_{D\times I} \frac{\partial v}{\partial t} w_1 + \int_{D\times I} \kappa(\cdot, \xi) \nabla v \cdot \nabla w_1 + \int_D v(\cdot, 0) w_2, \quad \text{and}$$
$$f(w; \xi) = \int_{D\times I} g(\cdot, \cdot, \xi) w_1 + \int_D u_0(\cdot, \xi) w_2.$$

- Assume $\tilde{\alpha} \leq \kappa(x, \xi) \leq \tilde{\beta}$.

Example 3 : diffusion equation on a random domain

$$-\Delta U(x,\xi) = g(x) \quad \text{for} \quad x \in D(\xi), \qquad U(x,\xi) = 0 \quad \text{for} \quad x \in \partial D(\xi)$$

- Assume $\phi(\cdot;\xi) : D_0 \rightarrow D(\xi)$ is a diffeomorphism from a deterministic domain D_0 to $D(\xi)$.
- Change of variable $u(x_0,\xi) = U(\phi(x_0,\xi),\xi)$, $x_0 \in D_0$.
- Bilinear form $a(u,w;\xi) = \int_{D_0} \nabla w(x_0) \cdot K(x_0,\xi) \cdot \nabla u(x_0)\, dx_0$, with $K = \nabla\phi\nabla\phi^T|\det(\nabla\phi)|$
- Linear form $f(w;\xi) = \int_{D_0} g_0(x_0,\xi) w(x_0) dx_0$, with $g_0(x_0,\xi) = g(\phi(x_0,\xi))|\det(\nabla\phi(x_0,\xi))|$
- Assumption on the diffeomorphism

$$\tilde{\alpha}\|\zeta\|_2 \leq \|\nabla\phi(x_0,\xi)\zeta\|_2 \leq \tilde{\beta}\|\zeta\|_2$$

- Approximation $u \in \mathcal{V} \subset H_0^1(D_0)$, $\mathcal{W} = \mathcal{V}$.

Operator equation and algebraic form

- 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 $A(\xi) : \mathcal{V} \to \mathcal{W}^*$ is an isomorphism such that

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

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

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

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

Regularity of the solution

- Regularity of the solution

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

If $\alpha(\xi) \geq \alpha_\star > 0$,

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

If $f \in L^p_\mu(\Xi; \mathcal{W}^*)$, then

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

- For $\alpha_\star = 0$ and/or $\gamma_\star = \infty$, see [Mugler-Starkloff 2011, Charrier 2012, Nouy-Soize 2014]
- From now on, assume

$$u \in L^2_\mu(\Xi; \mathcal{V})$$

Stochastic (or parametric) weak form

If $u(\xi)$ satisfies almost surely

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

then for all (measurable) functions $w : \Xi \to \mathcal{W}$

$$\mathbb{E}_\mu(\langle A(\xi)u(\xi), w(\xi)\rangle) = \mathbb{E}_\mu(\langle f(\xi), w(\xi)\rangle)$$

or

$$B(u, w) = F(w)$$

with

$$B(v, w) = \mathbb{E}_\mu(\langle A(\xi)v(\xi), w(\xi)\rangle) = \int_\Xi \langle A(y)v(y), w(y)\rangle \mu(dy)$$

$$F(w) = \mathbb{E}_\mu(\langle f(\xi), w(\xi)\rangle) = \int_\Xi \langle f(y), w(y)\rangle \mu(dy)$$

Weak formulation

Find $u \in X$ such that

$$B(u, w) = F(w) \quad \forall w \in Y \tag{1}$$

Stochatic (or parametric) weak form

Let

$$X = L^2_\mu(\Xi; \mathcal{V}), \quad Y = L^2_\mu(\Xi; \mathcal{W})$$

Under previous assumptions on $A(\xi)$, we deduce the following properties.

Properties of bilinear form $B : X \times Y \to \mathbb{R}$

- Continuous

$$\sup_{v \in X} \sup_{w \in Y} \frac{B(v,w)}{\|v\|_X \|w\|_Y} \leq \gamma_\star < \infty$$

- Weakly coercive

$$\inf_{v \in X} \sup_{w \in Y} \frac{B(v,w)}{\|v\|_X \|w\|_Y} \geq \alpha_\star > 0 \tag{2}$$

- $$\forall w \in X \setminus \{0\}, \quad \sup_{v \in Y} B(v,w) > 0 \tag{3}$$

Recall that (2) and (3) are satisfied if $B : X \times X \to \mathbb{R}$ is coercive :

$$\inf_{v \in X} \frac{B(v,v)}{\|v\|^2_X} \geq \alpha_\star > 0$$

Parametric (or stochastic) weak form

Theorem

If $F \in Y^* = L^2_\mu(\Xi; \mathcal{W}^*)$*, there exists a unique solution* $u \in X = L^2_\mu(\Xi; \mathcal{V})$ *to problem* (1) *and*

$$\|u\|_X \leq \frac{1}{\alpha_\star} \|F\|_{Y^*}$$

Galerkin methods

- Introduce approximation spaces

$$X_n \subset X$$

$$Y_n \subset Y$$

- Galerkin approximation defined by

$$u_n \in X_n \quad \text{such that} \quad B(u_n, w_n) = F(w_n) \quad \forall w_n \in Y_n$$

Galerkin methods

- Assume uniform stability of approximation spaces

$$\inf_{u_n \in X_n} \sup_{w_n \in Y_n} \frac{B(u_n, w_n)}{\|u_n\|_X \|w_n\|_Y} \geq \alpha_\star \tag{4}$$

  In particular, (4) is satisfied

  - if B is coercive and $X_n = Y_n$.
  - if $Y_n = \{w_n(\xi) = R_{\mathcal{W}}^{-1} A(\xi) v_n(\xi) : v_n \in X_n\}$ with $R_{\mathcal{W}}$ the Riesz map from $\mathcal{W}$ to $\mathcal{W}^*$.

- Quasi-optimality

$$\boxed{\|u - u_n\|_X \leq C \inf_{v \in X_n} \|u - v\|_X} \quad \text{with} \quad C = 1 + \frac{\gamma_\star}{\alpha_\star}$$

  The analysis of the best approximation error $\inf_{v \in X_n} \|u - v\|_X$ requires extra information on approximation spaces and the solution u (regularity).

- Convergence: For an increasing sequence of approximation spaces $X_n \subset X_{n+1}$ such that $\bigcup_{n \geq 1} X_n$ is dense in X, then $\|u - u_n\| \to 0$ when $n \to \infty$.
- Stability: For u_n and u_n Galerkin approximations of u and u', then

$$\|u_n - u'_n\|_X \leq \frac{\gamma_\star}{\alpha_\star} \|u - u'\|_X$$

Galerkin methods

- What are the classical choices for approximation spaces X_n ?
  - Projection-based model reduction

$$X_n = \mathcal{V}_n \otimes L^2_\mu(\Xi) = \{\sum_{i=1}^n v_i s_i(\xi) : s_i \in L^2_\mu(\Xi)\}$$

  - Stochastic Galerkin methods

$$X_n = \mathcal{V} \otimes \mathcal{S}_n = \{\sum_{j=1}^n u_j \psi_j(\xi) : u_j \in \mathcal{V}\}$$

- How does the best approximation $\inf_{v \in X_n} \|u - v\|_X$ behaves for these approximation spaces ?
- Can we characterize and compute optimal approximation spaces $\mathcal{V}_n$ and $\mathcal{S}_n$: relation with optimal low-rank approximation...

Outline

1 Functional framework for parametric and stochastic equations

2 Tensors

3 Low-rank approximation of order-two tensors

4 Computing low-rank approximations

5 Low-rank methods for parametric and stochastic equations

Tensor spaces

- Let $\mathcal{V}$ and $\mathcal{S}$ two vector spaces. The algebraic tensor space $\mathcal{V} \otimes \mathcal{S}$ is the set of elements of the form
$$\sum_{i=1}^{m} v_i \otimes s_i$$
with $v_i \in \mathcal{V}$, $s_i \in \mathcal{S}$, and $m \in \mathbb{N}$.
- A tensor Banach space is obtained by the completion of the algebraic tensor space $\mathcal{V} \otimes \mathcal{S}$ with respect to a norm $\|\cdot\|$:
$$\mathcal{V} \otimes_{\|\cdot\|} \mathcal{S} = \overline{\mathcal{V} \otimes \mathcal{S}}^{\|\cdot\|}$$

Examples of finite dimensional tensor spaces

- Matrices

$$a \in \mathbb{R}^{N\times M} = \mathbb{R}^N \otimes \mathbb{R}^M$$

$$a = \sum_{i=1}^{N} \sum_{j=1}^{M} a_{ij} e_i \otimes e_j$$

- Finite dimensional tensor spaces

$$\mathcal{V} \otimes \mathcal{S} = \overline{\mathcal{V} \otimes \mathcal{S}}^{\|\cdot\|}$$

Denoting $\{\phi_i\}_{i=1}^N$ a basis of $\mathcal{V}$ and $\{\psi_i\}_{i=1}^M$ a basis of $\mathcal{S}$, $u \in \mathcal{V} \otimes \mathcal{S}$ can be written

$$u = \sum_{i=1}^{N} \sum_{j=1}^{M} a_{ij} \phi_i \otimes \psi_j,$$

and identified with

$$a \in \mathbb{R}^N \otimes \mathbb{R}^M$$

Bochner spaces

- The Bochner space $L^p_\mu(\Xi;\mathcal{V})$ is 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

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

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

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

- Case $p=\infty$.

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

with equality if $\mathcal{V}$ is a Hilbert space or if μ is a discrete measure with finite support $\Xi_M = \{\xi_i\}_{i=1}^M : \mu = \sum_{\xi\in\Xi_M}\delta_{\xi_i}$, then $L^p_\mu(\Xi)\simeq\mathbb{R}^M$ and $L^p_\mu(\Xi;\mathcal{V}) = L^p_\mu(\Xi)\otimes\mathcal{V}\simeq\mathbb{R}^M\otimes\mathcal{V}$

Tensor norms

- We consider that $\mathcal{V}$ and $\mathcal{S}$ are normed spaces equipped with norms $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{S}}$.
- A necessary condition for a norm $\|\cdot\|$ on $\mathcal{V}\otimes\mathcal{S}$ is the continuity of the tensor product map $\otimes : \mathcal{V}\times\mathcal{S}\rightarrow\mathcal{V}\otimes\mathcal{S}$, that means the existence of C such that

$$\|v\otimes s\| \leq C\|v\|_{\mathcal{V}}\|s\|_{\mathcal{S}}$$

- A norm $\|\cdot\|$ is called a crossnorm if

$$\|v\otimes s\| = \|v\|_{\mathcal{V}}\|s\|_{\mathcal{S}}$$

  This property does not define a norm on the whole algebraic space $\mathcal{V}\otimes\mathcal{S}$.
- Norms $\|\cdot\|$ on $\mathcal{V}\otimes\mathcal{S}$ can be completely defined from the norms $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{S}}$. These are called *canonical* or *induced* norms.

Projective norm

- For $u \in \mathcal{V} \otimes \mathcal{S}$, the projective norm is defined by

$$\|u\|_\wedge = \inf \left\{ \sum_{i=1}^m \|v_i\|_\mathcal{V} \|s_i\|_\mathcal{S} : u = \sum_{i=1}^m v_i \otimes s_i \right\}$$

  where the infimum is taken over all representations of u.

- The projective norm is stronger than any norm $\|\cdot\|$ making continuous the tensor product map $\otimes : \mathcal{V} \times \mathcal{S} \to \mathcal{V} \otimes \mathcal{S}$, that means

$$\|\cdot\| \lesssim \|\cdot\|_\wedge$$

  so that

$$\mathcal{V} \otimes_{\|\cdot\|_\wedge} \mathcal{S} \subset \mathcal{V} \otimes_{\|\cdot\|} \mathcal{S}$$

Dual spaces

- For X a normed vector space, X' denotes the algebraic dual of X and X^* denotes the topological dual of X. We denote by $\|\cdot\|_X^*$ the dual norm to $\|\cdot\|_X$, defined for $\varphi \in X^*$ by $\|\varphi\|_X^* = \sup\{\varphi(x) : x \in X, \|x\|_X = 1\}$.
- For $\varphi \in \mathcal{V}'$ and $\psi \in \mathcal{S}'$, an element $\varphi \otimes \psi \in \mathcal{V}' \otimes \mathcal{S}'$ can be seen as a linear form on $\mathcal{V} \otimes \mathcal{S}$ via the definition

$$(\varphi \otimes \psi)(v \otimes s) = \varphi(v)\psi(s)$$

  so that

$$\mathcal{V}^* \otimes \mathcal{S}^* \subset \mathcal{V}' \otimes \mathcal{S}' \subset (\mathcal{V} \otimes \mathcal{S})'$$

- A norm $\|\cdot\|$ on $\mathcal{V} \otimes \mathcal{S}$ allows to define a dual space $(\mathcal{V} \otimes \mathcal{S})^*$ equipped with a dual norm denoted $\|\cdot\|^*$.
- If $\|\cdot\|$ is such that the tensor product map $\otimes : \mathcal{V}^* \times \mathcal{S}^* \to \mathcal{V}^* \otimes \mathcal{S}^*$ is continuous, that means

$$\|\varphi \otimes \psi\|^* \le C \|\varphi\|_{\mathcal{V}}^* \|\psi\|_{\mathcal{S}}^*$$

  for some constant C, then

$$\mathcal{V}^* \otimes \mathcal{S}^* \subset (\mathcal{V} \otimes \mathcal{S})^*$$

- A crossnorm $\|\cdot\|$ such that $\|\cdot\|^*$ is also a crossnorm is called a reasonable crossnorm. The projective norm is a reasonable crossnorm.

Injective norm

- For $u \in \mathcal{V} \otimes \mathcal{S}$, the injective norm is defined by

$$\|u\|_{\vee} = \sup\left\{(\varphi \otimes \psi)(u) : \varphi \in \mathcal{V}^*, \psi \in \mathcal{S}^*, \|\varphi\|_{\mathcal{V}}^* = \|\psi\|_{\mathcal{S}}^* = 1\right\}$$

- The injective norm is a reasonable crossnorm.
- The injective norm is weaker than any other norm $\|\cdot\|$ making the tensor product map $\otimes : \mathcal{V}^* \times \mathcal{S}^* \to \mathcal{V}^* \otimes \mathcal{S}^*$ continuous, that means

$$\|\cdot\| \gtrsim \|\cdot\|_{\vee} \quad (\|\cdot\|^* \lesssim \|\cdot\|_{\vee}^*)$$

so that

$$\mathcal{V} \otimes_{\|\cdot\|} \mathcal{S} \subset \mathcal{V} \otimes_{\|\cdot\|_{\vee}} \mathcal{S}$$

Hilbert tensor space

- Assume that $\mathcal{V}$ and $\mathcal{S}$ are Hilbert spaces equipped with inner products $\langle \cdot, \cdot \rangle_{\mathcal{V}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{S}}$.
- A canonical inner product $\langle \cdot, \cdot \rangle$ can be defined for $v, \tilde{v} \in \mathcal{V}$ and $s, \tilde{s} \in \mathcal{S}$ by

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

  and extended by linearity to $\mathcal{V} \otimes \mathcal{S}$.
- The associated norm $\| \cdot \|$ is a reasonable crossnorm.

Relation with operators

- Assume that $\mathcal{V}$ is a Hilbert space.
- $u = \sum_{i=1}^{m} v_i \otimes s_i \in \mathcal{V} \otimes \mathcal{S}$ can be identified with a linear operator from $\mathcal{V}$ to $\mathcal{S}$ such that for $v \in \mathcal{V}$

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

- The algebraic tensor space coincides with the set of finite rank operators

$$\mathcal{V} \otimes \mathcal{S} = \mathcal{F}(\mathcal{V}, \mathcal{S})$$

Relation with operators

- The injective norm $\|u\|_\vee$ coincides with the operator norm $\sup_{\|v\|_\mathcal{V}=1} \|u(v)\|_\mathcal{S}$, and

$$\overline{\mathcal{V}\otimes\mathcal{S}}^{\|\cdot\|_\vee} = \overline{\mathcal{F}(\mathcal{V},\mathcal{S})} = \mathcal{K}(\mathcal{V},\mathcal{S}),$$

  the set of compact operators.

- The tensor space equipped with the projective norm coincides with the set of nuclear operators

$$\overline{\mathcal{V}\otimes\mathcal{S}}^{\|\cdot\|_\wedge} = \mathcal{N}(\mathcal{V},\mathcal{S})$$

- If $\mathcal{S}$ is also a Hilbert space, the tensor space equipped with the canonical inner product norm $\|\cdot\|$ coincides with the space of Hilbert-Schmidt operators

$$\overline{\mathcal{V}\otimes\mathcal{S}}^{\|\cdot\|} = HS(\mathcal{V},\mathcal{S})$$

Singular value decomposition

- Assume $\mathcal{V}$ and $\mathcal{S}$ are Hilbert spaces.
- $u \in \mathcal{K}(\mathcal{V}, \mathcal{S})$ admits a singular value decomposition : there exist orthonormal systems $\{v_i\}$ in $\mathcal{V}$ and $\{s_i\} \in \mathcal{S}$, and a non increasing positive sequence $\{\sigma_i\}$ with $\sigma_i \searrow 0$ such that
$$u = \sum_{i=1}^{\infty} \sigma_i v_i \otimes s_i$$
which converges in the operator norm.
- Injective norm
$$\|u\|_{\vee} = \sigma_1$$
- Projective norm
$$\|u\|_{\wedge} = \sum_{i=1}^{\infty} \sigma_i$$
- The canonical inner product norm coincides with the Hilbert Schmidt norm
$$\|u\|_{HS}^2 = \sum_{i=1}^{\infty} \sigma_i^2$$

Coming back to Bochner spaces

- $L^1_\mu(\Xi; \mathcal{V}) = \overline{L^1_\mu(\Xi) \otimes \mathcal{V}}^{\|\cdot\|_1}$,

$$\|\cdot\|_1 = \|\cdot\|_\wedge$$

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

$$\|\cdot\|_\infty = \|\cdot\|_\vee$$

- $L^p_\mu(\Xi; \mathcal{V}) = \overline{L^p_\mu(\Xi) \otimes \mathcal{V}}^{\|\cdot\|_p}$ $(1 \leq p < \infty)$,

$$\|\cdot\|_\vee \leq \|\cdot\|_p \leq \|\cdot\|_\wedge$$

Outline

1 Functional framework for parametric and stochastic equations

2 Tensors

3 Low-rank approximation of order-two tensors

4 Computing low-rank approximations

5 Low-rank methods for parametric and stochastic equations

Low-rank approximation of 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 tensors with rank bounded by m

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

- Best approximation $u_m \in \mathcal{R}_m$ (provided it exists) of

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

with respect to $\|\cdot\|$ defined by

$$\|u - u_m\| = \min_{w \in \mathcal{R}_m} \|u - w\| \qquad (\star)$$

Minimal subspaces

- The minimal subspaces $U_1^{min}(w)$ and $U_2^{min}(w)$ of $w \in \mathcal{V} \otimes \mathcal{S}$ are the smallest subspaces in $\mathcal{V}$ and $\mathcal{S}$ respectively such that

$$w \in U_1^{min}(w) \otimes \mathcal{S} \quad \text{and} \quad \mathcal{V} \otimes U_2^{min}(w)$$

- For $w \in \mathcal{V} \otimes \mathcal{S}$

$$U_1^{min}(w) = \left\{ (I_d \otimes \psi)(w) : \psi \in \mathcal{S}' \right\}, \quad U_2^{min}(w) = \left\{ (\varphi \otimes I_d)(w) : \varphi \in \mathcal{V}' \right\}$$

- Rank of $w \in \mathcal{V} \otimes \mathcal{S}$

$$\boxed{rank(w) = dim(U_1^{min}(w)) = dim(U_2^{min}(w))}$$

Well-posedness of best approximation problem

- If $\|\cdot\| \gtrsim \|\cdot\|_\vee$, then
$$rank(\cdot) : \overline{\mathcal{V} \otimes \mathcal{S}}^{\|\cdot\|} \to \mathbb{R}$$
is weakly lower semi-continuous (w.l.s.c.) and therefore,
$$\mathcal{R}_m = \{w \in \mathcal{V} \otimes \mathcal{S} : rank(w) \leq m\}$$
is weakly closed.
- If $\|\cdot\| \gtrsim \|\cdot\|_\vee$ and $\overline{\mathcal{V} \otimes \mathcal{S}}^{\|\cdot\|}$ is reflexive, then a best approximation in $\mathcal{R}_m$ exists.
- If $\|\cdot\|$ is not stronger than $\|\cdot\|_\vee$ but the tensor space is an intersection of tensor spaces with such conditions on norms, well-posedness results can be obtained.

Low-rank approximation of order-two tensors: subspace point of view

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

$$\min_{u_m \in \mathcal{R}_m} \|u - u_m\| = \min_{dim(\mathcal{V}_m)=m} \min_{dim(\mathcal{S}_m)=m} \min_{u_m \in \mathcal{V}_m \otimes \mathcal{S}_m} \|u - u_m\|$$

or

$$\min_{u_m \in \mathcal{R}_m} \|u - u_m\| = \min_{dim(\mathcal{V}_m)=m} \min_{u_m \in \mathcal{V}_m \otimes \mathcal{S}} \|u - u_m\|$$

- That defines sequences of optimal subspaces $\mathcal{V}_m$ and $\mathcal{S}_m$ (w.r.t. the chosen norm $\|\cdot\|$). 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 $\mathcal{S}_m = span\{s_i\}_{i=1}^m$:

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

Low-rank approximation in $\mathcal{V} \otimes L^p_\mu(\Xi)$

- Natural (induced) norm

$$\|u\|_p = \left(\int_\Xi \|u(\xi)\|^p_{\mathcal{V}} \mu(d\xi) \right)^{1/p} \quad \text{for } p < \infty \quad \text{or} \quad \|u\|_\infty = \operatorname*{ess\,sup}_{\xi \in \Xi} \|u(\xi)\|_{\mathcal{V}}$$

- 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-m approximation solves

$$\min_{w \in \mathcal{R}_m} \|u - w\|_p = \min_{\dim(\mathcal{V}_m)=m} \min_{w \in \mathcal{V}_m \otimes L^p_\mu} \|u - w\|_p = \min_{dim(\mathcal{V}_m)=m} \|u - P_{\mathcal{V}_m} u\|_p$$

$$\text{with} \quad \|u(\xi) - P_{\mathcal{V}_m} u(\xi)\|_{\mathcal{V}} = \min_{v \in \mathcal{V}_m} \|u(\xi) - v\|_{\mathcal{V}}$$

- Relation with optimal projection-based model reduction

$$\min_{w \in \mathcal{R}_m} \|u - w\|_p = \min_{dim(\mathcal{V}_m)=m} \| \|u(\xi) - P_{\mathcal{V}_m} u(\xi)\|_{\mathcal{V}} \|_{L^p_\mu(\Xi)} := d_m^{(p)}(u)$$

- $d_m^{(p)}(u)$ is a linear width of the set of solutions $K = \{u(\xi) : \xi \in \Xi\} \subset \mathcal{V}$ that measures how well can be approximated by a m-dimensional space $\mathcal{V}_m$. It quantifies the ideal performance of a linear method.
  - For $p = \infty$, Kolmogorov m-width

$$d_m^{(\infty)}(u) := \min_{dim(\mathcal{V}_m)=m} \operatorname*{ess\,sup}_{\xi \in \Xi} \|u(\xi) - P_{\mathcal{V}_m} u(\xi)\|_{\mathcal{V}} \leq d_m(K)$$

  - For $p < \infty$, linear m-width for L^p_μ-optimality (measure-dependent)

$$d_m^{(p)}(u) := \min_{dim(\mathcal{V}_m)=m} \left(\int_\Xi \|u(\xi) - P_{\mathcal{V}_m} u(\xi)\|^p_{\mathcal{V}} \mu(d\xi) \right)^{1/p}$$

  - For $p = 2$, the best rank-m approximation is the truncated singular value decomposition of u and $d_m^{(2)}(u) = \left(\sum_{i>m} \sigma_i^2\right)^{1/2}$. Singular value decomposition also known as Karhunen-Loeve decomposition for μ a probability measure.

How to quantify optimal reduction methods ?

How fast m-widths go to zero with m ?

- Some general results in approximation theory (usually exploiting smoothness).
- Some finer results for particular cases.

Behaviour of m-widths

Consider the parametric model

$$-\nabla \cdot (a(x,\xi)\nabla u(\xi)) = f \quad \text{in } D \subset \mathbb{R}^d, \quad u(\xi) = 0 \quad \text{on } \partial D$$

$$0 < \alpha \leq a(x,\xi) \leq \gamma < \infty$$

- A general result.

$$K = u(\Xi) \subset H_0^1(D) = \mathcal{V}$$

If $f \in H^{s-1}(D)$ and $a(\cdot,\xi) \in C^s$, then $u(\xi) \in H^{s+1}$ and

$$d_m(K) \lesssim m^{-s/d}$$

Behaviour of m-widths

- Finer results taking into account the particular parametrization

$$a(x,\xi) = a_0(x) + \sum_{i=1}^{d} a_i(x)\xi_i, \quad \xi_i \in (-1,1)$$

  - $d < \infty$: Exponential convergence of $d_m(K)$. Deterioration of the rate with d.
  - $d = \infty$: If $(\|a_i\|_\infty)_{i\geq 1} \in \ell_p$ with $p < 1$, then [Cohen-DeVore-Schwab 2010]

$$d_m(K) \lesssim m^{-1/p+1}$$

- Towards general results [DeVore et al 2014]. Considering

$$\mathcal{A} = \{a(\cdot,\xi) : \xi \in \Xi\} \subset C(D),$$

then

$$d_m(K) \lesssim d_m(\mathcal{A})$$

Behaviour of m-widths: relation with best-m term approximation

- Bounds of m-widths can be obtained from best m-term approximations.
- Let $\{\psi_\alpha\}_{\alpha\in\Lambda}$ be any set of functions. For $\Lambda_m \subset \Lambda$, let $\mathcal{S}_{\Lambda_m} = \mathrm{span}\{\psi_\alpha\}_{\alpha\in\Lambda_m}$.
- We have

$$d_m^{(p)}(u) \leq \inf_{\#\Lambda_m = m} \inf_{w\in\mathcal{V}\otimes\mathcal{S}_{\Lambda_m}} \|u-w\|_{L^p_\mu(\Xi;\mathcal{V})}$$

  that means

$$d_m^{(p)}(u) \leq \|u-u_{\Lambda_m}\|_{L^p_\mu(\Xi;\mathcal{V})}$$

  for any m-dimensional subspace $\mathcal{S}_{\Lambda_m}$ and any approximations u_{Λ_m} in $\mathcal{V}\otimes\mathcal{S}_{\Lambda_m}$.
- Convergence results for $\|u-u_{\Lambda_m}\|_{L^p_\mu(\Xi;\mathcal{V})}$ then provide estimates for m-width $d_m^{(p)}(u)$.

Outline

1. Functional framework for parametric and stochastic equations
2. Tensors
3. Low-rank approximation of order-two tensors
4. Computing low-rank approximations
5. Low-rank methods for parametric and stochastic equations

Optimal low-rank approximation in the general case

- In general, best rank-m approximation (provided it exists) can be defined w.r.t. to a certain distance to the solution

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

- If

$$\mathcal{E}(u, w) \sim \|u - w\|$$

then

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

- $\mathcal{R}_m$ is a manifold (not linear space nor convex set) : nonlinear approximation problem.

Computing low-rank approximation in the general case

- In the Hilbert case and if $\mathcal{E}(u,w) = \|u-w\|_{HS}$ (induced canonical norm), then truncated SVD provides optimal low-rank approximations.
- Direct optimization in $\mathcal{R}_m$ using
  - Alternating minimization algorithms

$$\tilde{u}_m^{(k)} = \arg \min_{w \in \mathcal{V} \otimes \mathcal{S}_m^{(k-1)}} \mathcal{E}(u,w), \quad \mathcal{V}_m^{(k)} = U_1^{min}(\tilde{u}_m^{(k)})$$

$$u_m^{(k)} = \arg \min_{w \in \mathcal{V}_m^{(k)} \otimes \mathcal{S}} \mathcal{E}(u,w), \quad \mathcal{S}_m^{(k)} = U_2^{min}(u_m^{(k)})$$

  - other algorithms on manifolds

Computing low-rank approximation in the general case

- Except for the Hilbert case with induced canonical norm $\mathcal{E}(u,w) = \|u-w\|_{HS}$,
  - Optimal subspaces are not necessarily nested

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

  - No notion of decomposition

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

- Suboptimal approximation using constructive algorithms : greedy construction of approximation or subspaces
  - Reduced Basis method (greedy algorithms) and Generalized 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 $\mathcal{R}_1$:

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

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

- 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 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 $\{v_1, \ldots, v_m, \ldots\}$.

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

Outline

1. Functional framework for parametric and stochastic equations
2. Tensors
3. Low-rank approximation of order-two tensors
4. Computing low-rank approximations
5. Low-rank methods for parametric and stochastic equations

Parametric and stochastic models

$$u(\xi) \in \mathcal{V}, \quad A(\xi)u(\xi) = f(\xi)$$

with $A(\xi) : \mathcal{V} \to \mathcal{W}^*$ and $f(\xi) \in \mathcal{W}^*$

Tensor structured equations

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

$$A(\xi) = \sum_{k=1}^{R} \lambda_k(\xi) A_k, \quad f(\xi) = \sum_{k=1}^{L} \eta_k(\xi) f_k$$

- If no such low-rank representation of operator and right-hand-side (or if R and L are high), preliminary approximation (e.g. using interpolation)

Example

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

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

Tensor-structured equations for Galerkin approximation

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 for Galerkin approximation

- $\lambda : \Xi \to \mathbb{R}$ 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 $\{\Phi_i\}_{i=1}^M$ and $\{\Psi_i\}_{i=1}^M$ bases of $\mathcal{S}$ and $\widetilde{\mathcal{S}}$, algebraic representation of Λ:

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

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

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

- Tensor structured equation in algebraic form

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

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 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

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

Remark: another residual-based error

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

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

- 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-Nouy-Legrain 2014]
- 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 [Billaud-Nouy-Zahm 2014]

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

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

  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)

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

- First idea: Compute K samples of the solution, extract an optimal reduced basis for the samples using empirical KL, project the initial model on this basis (POD-like approach)

Low-rank approximation using sampling-based approach

- Second idea : Residual based approach

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

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

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

- Best rank-m approximation defined by

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

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

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

Low-rank approximation using sampling-based approach

- Set of equations

$$A(\xi)u(\xi) = f(\xi), \quad \xi \in \Xi_K \qquad (\square)$$

with

$$A(\xi) = \sum_{i=1}^{R} A_i \lambda_i(\xi), \quad f(\xi) = \sum_{i=1}^{L} f_i \eta_i(\xi)$$

- $(\square)$ identified with

$$\mathbf{Bu} = \mathbf{F}$$

with

$$\mathbf{B} = \sum_{i=1}^{R} A_i \otimes \mathbf{\Lambda}_i, \quad \mathbf{\Lambda}_i = diag(\lambda_i(\xi^1), \ldots, \lambda_i(\xi^K)) \in \mathbb{R}^{K \times K}$$

$$\mathbf{F} = \sum_{i=1}^{L} f_i \boldsymbol{\eta}_i, \quad \boldsymbol{\eta}_i = (\eta_i(\xi^1), \ldots, \eta_i(\xi^K))^T \in \mathbb{R}^K$$

Computing optimal low-rank approximation

- We have seen different ways of defining a low-rank approximation u_m by minimization a certain distance $\mathcal{E}(u, u_m)$ to the solution:

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

- $\mathcal{R}_m$ is a manifold (not linear space nor convex set) : nonlinear approximation problem.
  - Optimization in $\mathcal{R}_m$ using alternating direction algorithms or other optimization algorithms on manifolds.
  - Suboptimal approximation using constructive algorithms : greedy construction of approximation or subspaces, e.g. Proper Generalized Decomposition

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

$$\mathcal{E}(u, \sum_{i=1}^m v_i \otimes s_i^m) = \min_{v_m \in \mathcal{V}} \min_{(s_1, \ldots, s_m) \in \mathcal{S}^m} \mathcal{E}(u, \sum_{i=1}^m v_i \otimes s_i) \quad (5)$$

- Alternating minimization algorithm for solving (5): solve successively

$$\min_{v_m \in \mathcal{V}} \mathcal{E}(u, \sum_{i=1}^{m} v_i \otimes s_i)^2, \tag{6}$$

$$\min_{(s_1,\ldots,s_m)\in \mathcal{S}^m} \mathcal{E}(u, \sum_{i=1}^{m} v_i \otimes s_i)^2 \tag{7}$$

- Consider a symmetric problem, and let

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

- Solution of (6) (non parametric problem):

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

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 (7) (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

$$\boxed{\mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{A}_m(\xi)\mathbf{s}(\xi)) = \mathbb{E}_\mu(\mathbf{t}(\xi)^T\mathbf{f}_m(\xi)) \quad \forall \mathbf{t}\in(\mathcal{S})^m} \tag{8}$$

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 (8) is the stochastic Galerkin approximation of the solution of

$$\boxed{\mathbf{A}_m(\xi)\mathbf{s}(\xi) = \mathbf{f}_m(\xi)} \tag{9}$$

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

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

- (8) 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 reduce the cost (sparsity, low-rank...).
- (9) can be solved with sampling-based approaches (interpolation, regularized least-squares...)

Example: stochastic Groundwater flow equation (MOMAS/Couplex)

Groundwater flow equation (hydraulic head u)

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

+ boundary conditions

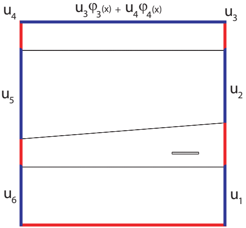
Geological layers with uncertain properties

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,
$\boxed{\Xi = (-1, 1)^{10}}$, uniform probability P_ξ

Uncertain BCs

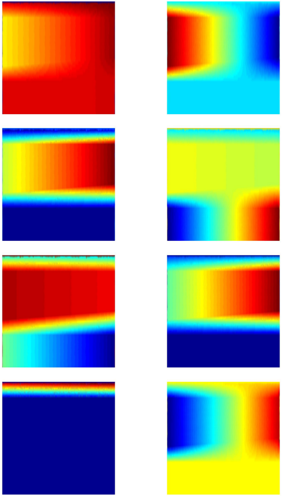


Neumann homogeneous
Dirichlet

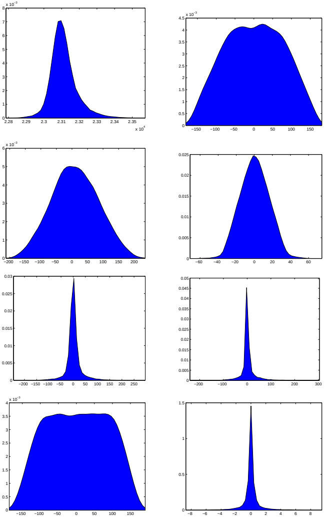
	Law
u_1	$U(288, 290)$
u_2	$U(305, 315)$
u_3	$U(330, 350)$
u_4	$U(170, 190)$
u_5	$U(195, 205)$
u_6	$U(285, 287)$

First modes with the greedy construction of the approximation

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

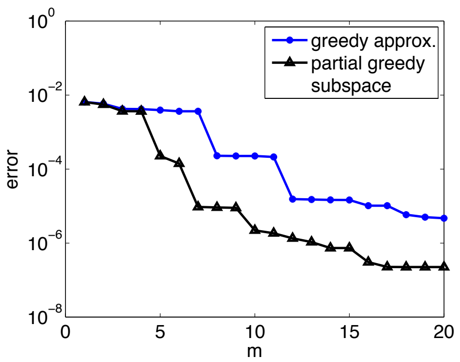


Stochastic modes $\{s_1, \dots, s_8\}$: pdf



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

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

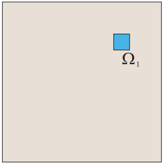
  Necessary (but not sufficient) condition of optimality

$$\langle Bu_m - F, \delta w \rangle = 0 \quad \forall \delta w \in T_{u_m}\mathcal{M}_m \tag{10}$$

  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 (10).
- Alternating direction algorithms yields problems with the same structure as previously.
- Heuristic approach. No theoretical results except for particular cases.

Application to an advection-diffusion-reaction equation

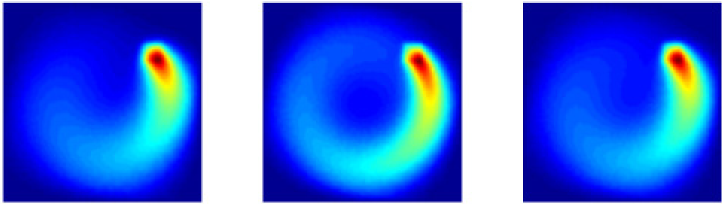
- $\partial_t u - a_1 \Delta u + a_2 c \cdot \nabla u + a_3 u = a_4 I_{\Omega_1}$ *on* $\Omega \times (0, T)$
- $u = 0$ *on* $\Omega \times \{0\}$
- $u = 0$ *on* $\partial\Omega \times (0, T)$



Uncertain parameters

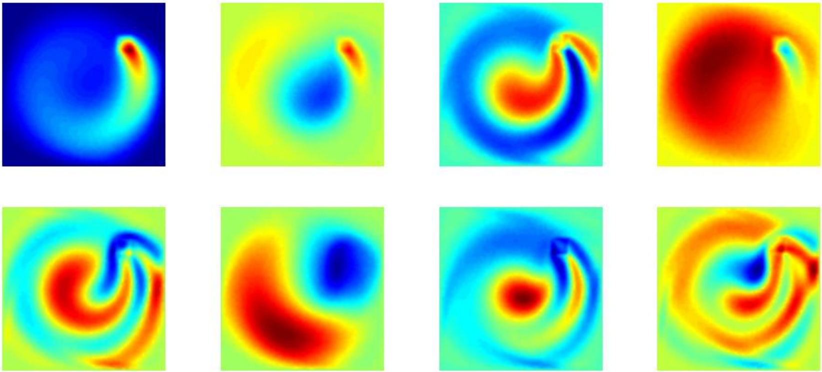
$$a_i(\xi) = \mu_{a_i}(1 + 0.2\xi_i), \quad \xi_i \in U(-1,1), \quad \boxed{\Xi = (-1,1)^4}$$

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



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

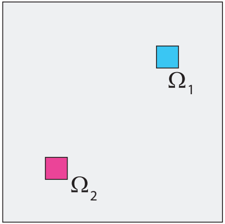
8 first modes of the decomposition $\{v_1(x,t) ... v_8(x,t)\}$



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

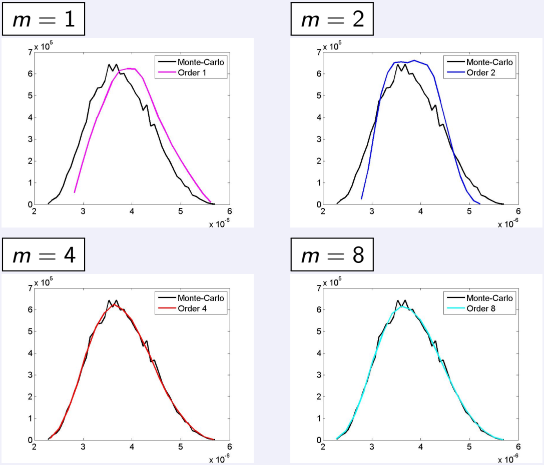
Quantity of interest

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



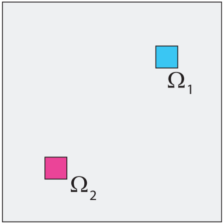
$$s_m(\xi) = \int_0^T \int_{\Omega_2} u_m(x,t,\xi)\, dxdt$$

Probability density function of $s_m(\xi)$



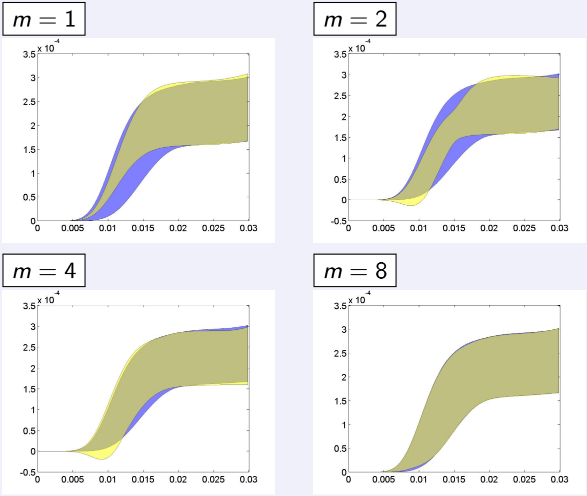
Quantity of interest

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



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

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



In summary

- Linear methods for order reduction yield an approximation of the form

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

  with $v_i \in \mathcal{V}$ and $s_i \in L^p_\mu(\Xi)$, which is an element of rank m in $\mathcal{V} \otimes L^p_\mu(\Xi)$
- Optimal linear order reduction methods are related with optimal low-rank approximation.
- Efficient solution methods exploiting low-rank formats
- Extension of these ideas to higher order tensor spaces ? Application to high-dimensional approximation...

References I

- **PROJECTION-BASED MODEL REDUCTION**

  A. Nouy. Low-rank tensor approximation methods and Proper Generalized Decompositions. Book Chapter. Preprint.

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

  A. Quarteroni, G. Rozza, and A. Manzoni. Certified reduced basis approximation for parametrized partial differential equations and applications. *Journal of Mathematics in Industry*, 1(1), 2011.

  A. Nouy. A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1603–1626, 2010.

- **STOCHASTIC SPECTRAL METHODS, GALERKIN METHODS**

  O. P. Le Maitre and O. M. Knio. *Spectral Methods for Uncertainty Quantification With Applications to Computational Fluid Dynamics*. Scientific Computation. 2010.

  A. Nouy. Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations. *Archives of Computational Methods in Engineering*, 16(3):251–285, 2009.

References II

- **TENSORS, OPERATORS**

A. Defant and K. Floret.
Tensor norms and operator ideals.
North-Holland, Amsterdam New York, 1993.

W. Hackbusch.
Tensor spaces and numerical tensor calculus, volume 42 of *Springer series in computational mathematics*.
Springer, Heidelberg, 2012.