Uncertainty Quantification and Model Reduction

Anthony Nouy

Ecole Centrale Nantes, GeM
Nantes, France

Outline

Uncertainties

Many models contain uncertainties that may be

- aleatoric, random and not reducible,
- epistemic, due to lack of knowledge, reducible by adding information.

Probability theory as a mathematical theory for modeling

- aleatoric uncertainty: frequentist point of view,
- and also epistemic uncertainty: probability arises from information theory (Cox's theorem, maximum of entropy), Bayesian point of view

Uncertainty quantification

Uncertainty quantification has become an essential path in science and engineering.

- Predictive modelling and simulation under uncertainty
- Interplay between numerical analysis, scientific computing, probability and statistics

Objectives:

- Assess confidence in model predictions
- Comprehension and selection of models (exploration of model predictions over a range of uncertainty, sensitivity analysis)
- Robust optimization, control

Uncertainty quantification

Requires an accurate description of input uncertainties. Probabilistic model must be constructed objectively based on

- prior information,
- and/or available data.

As an example, consider a diffusion model

$$-\nabla \cdot (a \nabla u) = f$$

with uncertain diffusion coefficient a.

- Prior information and possible use of maximum of entropy principle:
  - positive definite matrix (second law of thermodynamics) with possible symmetries (isotropy, orthotropy...), bounds, ...
  - stationarity, correlation structure...
- If data available, statistical inference or bayesian update of prior models. Data usually given as indirect observations with uncertainties.

Uncertainty quantification

Stochastic model

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

where ξ are random parameters 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 $s(\xi)$, estimate μ.

Computational approaches for forward problems

Different computational approaches:

- Direct integration : compute statistics using integration, e.g. with Monte-Carlo methods, Quasi-Monte-Carlo methods or other quadrature methods...
- Moment methods: equations for statistical moments, e.g. using perturbation methods
- Probability distributions : equations for probability densities, e.g. Kolmogorov equations
- Direct approximation : compute an approximation of $u(\xi)$ and then evaluate (quickly) output variables of interest, observables, or objective function. Approximation known as reduced order model, metamodel, surrogate model...

Complexity issues for high-dimensional problems

Complexity issues for direct approximation

- Complex numerical models

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

$$\dim(V) \gg 1$$

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

- Approximation of multivariate functions

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

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

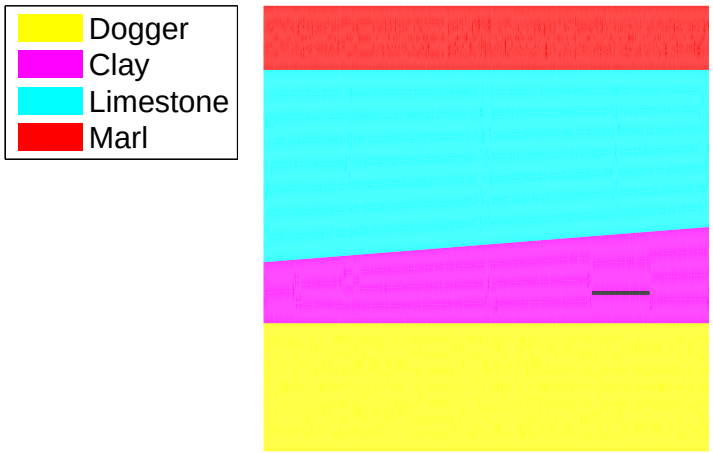
  - Classical approaches suffer from the curse of dimensionality
  - Remedy: adapted bases, structured approximations

Outline

An introductive example

Groundwater flow equation.

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



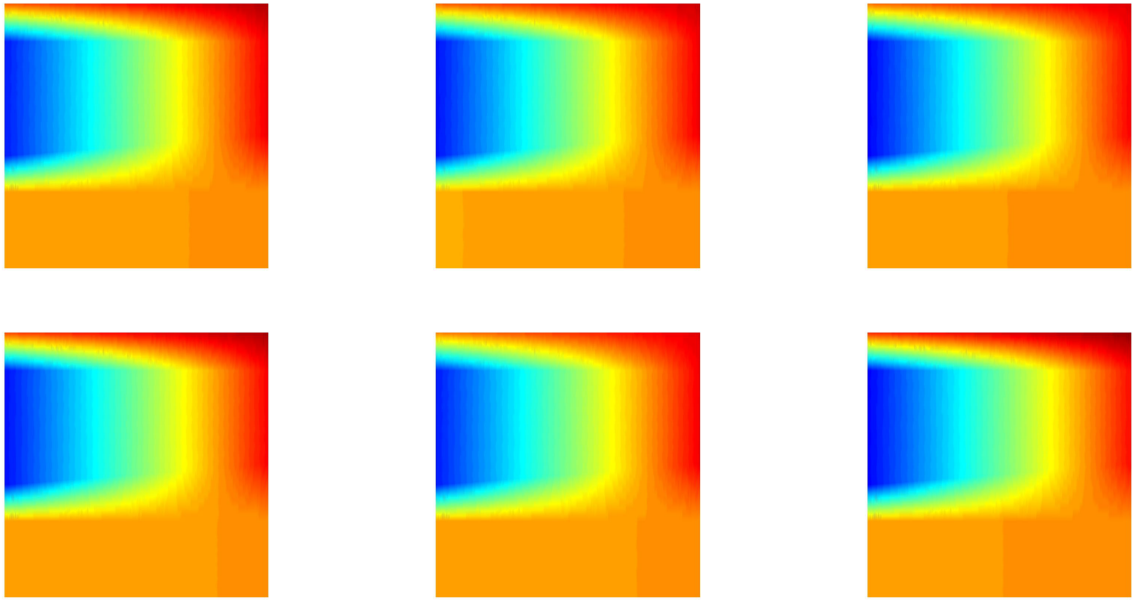
$$a(x, \xi) = \sum_{k=1}^{4} \xi_k I_{D_k}(x)$$

Layer	Random diffusion coefficient
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	$\xi_4 \sim LU(10^{-5}, 10^{-4})$

Problem requiring fine discretization :

$$u(\xi) \in V, \quad \dim(V) \gg 1$$

- Samples $\{u(\xi^i)\}_{1\leq i\leq m}$ of the solution $u(\xi)$.



- Low-dimensionality of the set of solutions $u(\xi)$ that should be well approximated on low-dimensional subsets.
- Interpolation as a first idea

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

Worst case error

$$\epsilon_m^{(\infty)} = \sup_{\xi} \|u(\xi) - u_m(\xi)\|$$

m	1	3	6	10
$\epsilon_m^{(\infty)}/\epsilon_0^{(\infty)}$	2.10^{-2}	8.10^{-3}	4.10^{-4}	1.10^{-6}

Model order reduction for parametric problems

$$u : \Xi \to V \quad \text{such that} \quad A(u(\xi); \xi) = f(\xi), \quad \xi \in \Xi$$

- Linear methods: approximation in a low-dimensional subspace

$$V_m = span(v_1, \ldots, v_m) \subset V$$

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

- Nonlinear methods: approximation in a low-dimensional manifold

$$V_m \subset V$$

$$u_m(\xi) = R(s_1(\xi), \ldots, s_m(\xi)) \quad \text{with} \quad R : \mathbb{R}^m \to V_m \text{ (parametrization of } V_m)$$

How to quantify optimal reduction methods ?

- How to measure the quality of a space (or manifold) V_m ?

$$\text{for } \xi \in \Xi, \quad d(u(\xi), V_m) = \inf_{v \in V_m} \|u(\xi) - v\|$$

  - Worst case error ($L^\infty(\Xi)$)

$$\sup_{\xi \in \Xi} d(u(\xi), V_m)$$

  - Averaged error ($L^p_\mu(\Xi)$)

$$\left(\int_\Xi d(u(\xi), V_m)^p \mu(d\xi) \right)^{1/p}$$

  - ...

- The ideal performance of a reduction method can be quantified by measuring the "width" of the solution manifold

$$K = \{u(\xi) : \xi \in \Xi\} \subset V$$

Linear widths (resp. nonlinear widths) measure how well K can be approximated by a linear space (resp. manifold)

How to quantify optimal reduction methods ?

- Kolmogorov width (linear width for $L^\infty(\Xi)$-optimality)

$$d_m(K) = \inf_{\dim(V_m)=m} \sup_{\xi\in\Xi} d(u(\xi), V_m)$$

Provides a benchmark for optimal reduced basis methods : if $d_m(K) \leq \epsilon$, there exists a m-dimensional space such that

$$d(u(\xi), V_m) \leq \epsilon \quad \text{for all } \xi$$

- Linear width for $L^p_\mu(\Xi)$-optimality

$$\sigma_m^{(p)}(u) = \inf_{\dim(V_m)=m} \left(\int_\Xi d(u(\xi), V_m)^p \mu(d\xi) \right)^{1/p}$$

The case $p = 2$ corresponds to singular value decomposition (Karhunen-Loeve decomposition) of u.

- Non-linear width measures how well K can be approximated by a manifold. Devore-Howard-Michelli width defined by

$$\delta_m(K) = \inf_{R,C} \sup_{\xi\in\Xi} \|u(\xi) - R(C(u(\xi)))\|$$

where the infimum is taken over all continuous maps $C : K \to \mathbb{R}^m$ and $R : \mathbb{R}^m \to V$.

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

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

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

- A general result.

$$K \subset H_0^1(D) = V$$

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

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

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

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

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

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

then

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

How to construct (quasi-)optimal reduced spaces ?

- $u(\xi)$ is not directly accessible but solves

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

- Approximation $u_m(\xi)$ defined through Galerkin-type projections, such as

$$u_m(\xi) = \arg \min_{v \in V_m} \|A(v;\xi) - f\|$$

Under continuity and stability assumptions,

$$\|u(\xi) - u_m(\xi)\|_V \leq c \inf_{v \in V_m} \|u(\xi) - v\|_V \quad \text{(quasi-optimality)}$$

- Constructing optimal spaces is usually intractable $\longrightarrow$ suboptimal nested spaces (constructive approach)
  - Reduced Basis Method for $L^\infty(\Xi)$

$$V_m = span\{u(\xi^1), \ldots, u(\xi^m)\} \quad \text{with adaptive sampling}$$

  - Proper Generalized Decomposition for $L^2_\mu(\Xi)$

$$\inf_{\dim(V_m)=m} \int_\Xi \|A(u_m(\xi);\xi) - f\|^2 \mu(d\xi) \quad \text{subject to} \quad V_m \supset V_{m-1}$$

In summary

For the model reduction of parametric problems, we are able (in some situations) to

- quantify optimal model reduction
- provide a priori estimates for the performance
- propose algorithms that achieves quasi-optimality

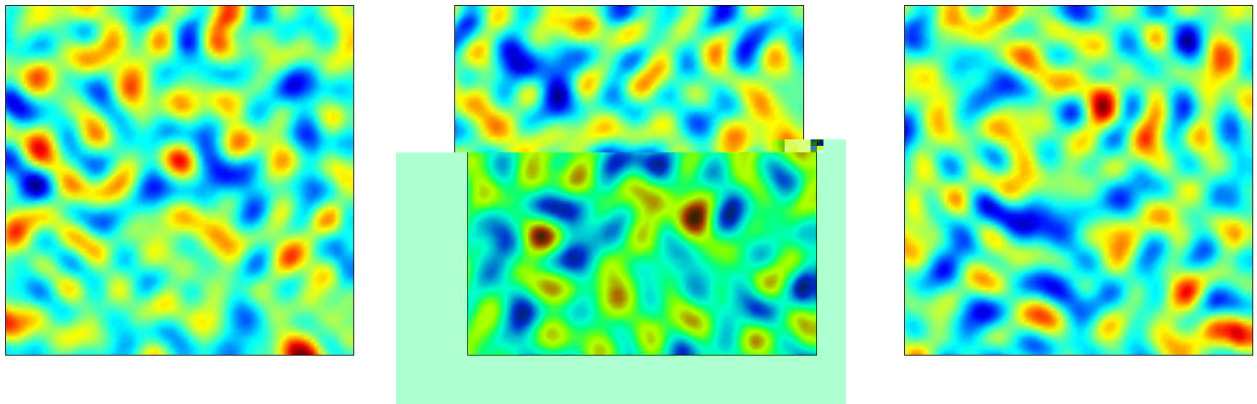
Outline

Examples of high dimensional problems in stochastic analyses

- Complex models with many uncertain parameters (geometry, source terms, constitutive laws)
- Random media with spatially correlated random fields

$$-\nabla \cdot (a \nabla u) = f \quad + \quad b.c.$$

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



Examples of high dimensional problems in stochastic analyses

- Stochastic differential equations:

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

  - Discretization of Wiener process $W_t \approx \sum_{k=1}^{d} \varphi_k(t)\xi_k$, and

$$X_t \approx X_t(\xi_1, \dots, \xi_d)$$

  - n-dimensional PDEs for probability density function

$$u(x_1, \dots, x_n, t)$$

    of X_t (Kolmogorov equation), or for functionals of X_t (Feynman-Kac formula)...

High-dimensional approximation

- Many problems requires the approximation (or integration) of high dimensional functions

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

- Classical discretization methods yields high-dimensional parametrizations

$$u(x_1, \ldots, x_d) \approx \sum_{i_1=1}^{n} \ldots \sum_{i_d=1}^{n} a_{i_1 \ldots i_d} \varphi_{i_1}^1(x_1) \ldots \varphi_{i_d}^d(x_d), \quad a \in \mathbb{R}^{n^d}$$

  - Number of parameters

$$N = n^d$$

  - Assuming $u \in C^s((0,1)^d)$, accuracy

$$\epsilon = O(n^{-s}) = O(N^{-s/d})$$

    and the number of parameters to achieve accuracy ϵ is

$$N = O(\epsilon^{-d/s}) \quad \text{(Curse of dimensionality)}$$

Curse of dimensionality - tractability

- Curse of dimensionality related to computational intractability in high dimension.
  - Quantitive definition through information-based complexity analysis
  - Tractability depends on the measure of precision and the available information.

- $N(\epsilon, d)$ being the number of linear informations to obtain a precision ϵ, intractability (curse of dimensionality) when

$$\lim_{\epsilon^{-1}+d \to \infty} \frac{\log N(\epsilon, d)}{\epsilon^{-1} + d} > 0$$

See [Novak,Wozniakowski]

Curse of dimensionality - tractability

- Weak tractability

$$\lim_{\epsilon^{-1}+d\to\infty} \frac{\log N(\epsilon, d)}{\epsilon^{-1} + d} = 0$$

- Polynomial tractability

$$N(\epsilon, d) \leq C\epsilon^{-p} d^q \quad \text{for all } \epsilon \text{ and } d$$

- Strong polynomial tractability

$$N(\epsilon, d) \leq C\epsilon^{-p} \quad \text{for all } \epsilon$$

Intractability for the approximation of smooth functions

Consider the set of functions

$$F_d = \left\{ u \in C^\infty((0,1)^d) : \sup_\alpha \|D^\alpha u\|_\infty < \infty \right\}$$

- Minimal approximation error using linear informations:

$$\epsilon(N,d) = \inf_{A_N} \sup_{u \in F_d} \|u - A_N(u)\|_\infty$$

  where the infimum is taken over all algorithms A_N using N linear informations on u and providing an approximation $A_N(u)$.
- Optimal rate of convergence is infinite: for arbitrary large s

$$\epsilon(N,d) = O(N^{-s}) \quad \text{as } N \to \infty$$

$$N(\epsilon,d) = O(\epsilon^{-1/s}) \quad \text{as } \epsilon \to 0$$

- But what about the constants in O ? [Novak 2009] proves

$$N(\epsilon,d) \geq 2^{\lfloor \frac{d}{2} \rfloor}$$

Curse of dimensionality !

Remedies: structured approximations

Order reduction methods must exploit specific structures (application dependent)

- Smoothness (anisotropy)
- Low effective dimensionality, e.g.

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

- Low-order interactions, e.g.

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

- Sparsity (relatively to a basis or frame)

$$u(x_1, \ldots, x_d) = \sum_{\alpha \in \Lambda} a_\alpha \psi_\alpha(x_1, \ldots, x_d) \approx \sum_{\alpha \in \Lambda_n \subset \Lambda} a_\alpha \psi_\alpha(x_1, \ldots, x_d)$$

- Low-rank structure

$$u(x_1, \ldots, x_d) \approx \sum_{i=1}^{r} u_{1,i}(x_1) \ldots u_{d,i}(x_d)$$

- ...

Outline

Sparse approximation

- We want to approximate a function $u \in L^p_\mu(\Xi; V)$. Consider a set of functions $\{\psi_\alpha; \alpha \in \Lambda\}$ (called a dictionary) whose span is dense in $L^p_\mu(\Xi)$, so that

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

  Typically, for multivariate functions

$$\psi_\alpha(\xi) = \psi^1_{\alpha_1}(\xi_1) \dots \psi^d_{\alpha_d}(\xi_d), \quad \Lambda = \{\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d; \sum_{i=1}^d \alpha_i < \infty\}$$

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

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

- Sparse approximations can be obtained though interpolation, regression, Galerkin projections.

Best n-term approximation

The best n-term approximation in L^p norm is the solution of

$$\min_{\#\Lambda_n = n} \inf_{u_n \in X_{\Lambda_n}} \|u - u_n\|_p := \sigma_n^{(p)}(u)$$

where

$$X_{\Lambda_n} = \left\{ v = \sum_{\alpha \in \Lambda_n} v_\alpha \psi_\alpha ; v_\alpha \in V \right\}$$

Best n-term approximation

Questions:

- How fast $\sigma_n^{(p)}(u)$ converges to zero ?
  - If $(\|u_\alpha\|_V)_\alpha \in \ell^q$ and $(\psi_\alpha)_\alpha$ are orthonormal in L^2, then there exists Λ_n such that

$$\sigma_n^{(2)} \leq \min_{u_n \in X_{\Lambda_n}} \|u - u_n\|_2 \leq C n^{-s} \quad \text{with} \quad s = 1/q - 1/2$$

  - See  [Cohen-DeVore-Schwab 2010,Chkifa-Cohen-Schwab 2014] for a proof of $(\|u_\alpha\|_V)_\alpha \in \ell^p$ for a large class of parametric problems and for polynomial approximation. Results are working for infinitely many random variables.
- Selection of basis (or frame) suitable for a class of functions ?
- Best n-term approximation is a combinatorial problem. How to construct Λ_n in practice ?

Sparse approximation in practice

- A priori definition of the index set Λ_n, e.g. based on the properties of the class of functions we want to approximate (low-order interactions, anisotropy...). A priori error estimation can be useful.
  - $\Lambda_n = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \alpha_i \leq p\}$
  - $\Lambda_n = \{\alpha \in \mathbb{N}^d; \sum_{i=1}^d \gamma_i \alpha_i \leq p\}$ (anisotropic)
  - $\Lambda_n = \{\alpha \in \mathbb{N}^d; f(\alpha) \leq p\}$.
- Adaptive construction based on a posteriori error estimation, e.g. by greedy algorithms:

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

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

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

Outline

Approximation of rank one functions

Consider the function $u : [0,1]^d \to \mathbb{R}$

$$u(x_1, \ldots, x_d) = u_1(x_1) \ldots u_d(x_d), \quad u_k \in C^s(0,1)$$

- Approximation of the factors

$$u_k(x_k) \approx \sum_{i=1}^{n} c_{k,i} \varphi_i(x_k) \quad \text{with error} \sim n^{-s}$$

- Number of parameters

$$N = dn$$

- Number of parameters to achieve accuracy ϵ

$$N(\epsilon, d) = O(d^{1+1/s} \epsilon^{-1/s})$$

Complexity of approximating rank-one tensors

- Consider the set of functions [Bachmayr 2013,Novak 2004]

$$F = \left\{ u(x_1, \ldots, x_d) = \prod_{i=1}^{d} u_i(x_i) : \|u_i\|_\infty \leq 1, \ \|u_i^{(s)}\|_\infty \leq M \right\}$$

  - If $M \geq 2^s s!$,

$$N(\epsilon, d) \geq 2^d \quad \text{for all } \epsilon < 1 \quad \text{(curse of dimensionality)}$$

  - Polynomial tractability (algorithm available) for the sets of functions

$$F^{x^*} = \{u \in F : u(x^*) \neq 0 \quad \text{for a known } x^*\}$$

$$F^V = \{u \in F : u \neq 0 \text{ on a box of measure greater than } V\}$$

Approximation with low-rank tensors (canonical format)

- Consider that function $u : [0,1]^d \to \mathbb{R}$ can be approximated with accuracy ϵ by an approximation with canonical rank $r(\epsilon)$:

$$u(x_1, \ldots, x_d) \approx \sum_{i=1}^{r(\epsilon)} u_{1,i}(x_1) \ldots u_{d,i}(x_d), \quad u_{k,i} \in C^s(0,1)$$

  - Approximation of each factor $u_{k,i}$ using n parameters with accuracy $\sim n^{-s}$
  - Number of parameters $N = r(\epsilon)nd$
  - Number of parameters to achieve accuracy ϵ

$$N(\epsilon, d) = O(r(\epsilon)^{1/s} d^{1+1/s} \epsilon^{-1/s})$$

- Do we beat the curse of dimensionality ?
  - What about $r(\epsilon)$ with respect to d ? some positive results...
  - Which information on u and which algorithm ?

Low-rank approximations

- Low-rank tensor subsets

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

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

- Canonical format has bad properties (non existence of best approximations, unstable format)
- Subspace based formats (Tucker or tree-based Tucker) have nice topological and geometrical properties (stability of best approximation, differentiable Banach manifolds)  [Falco-Hackbusch-Nouy 2014]
  - Existence of best approximations
  - Optimization on manifolds
  - Dynamical systems on manifolds (extension of Dirac-Frenkel theory to Banach tensor manifolds)

Operator equation in tensor format

$$A(u) = b \quad \text{with} \quad u \in X_{\|\cdot\|}, \quad X = \bigotimes_{\mu=1}^{d} V_\mu$$

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

Example: Laplace operator

$$A = \partial^2_{x_1} \otimes I \otimes \ldots \otimes I + \ldots + I \otimes \ldots \otimes I \otimes \partial^2_{x_d}$$

$$X = H^1_0(\Omega_1) \otimes \cdots \otimes H^1_0(\Omega_d)$$

Example: Parameter dependent operators (with affine decomposition)

$$A(\xi) = \sum_{i=1}^{m} A_i \lambda_i(\xi) \quad \text{or} \quad A = \sum_{i=1}^{m} A_i \otimes \Lambda_i$$

$$X = V \otimes L^p_\mu(\Xi)$$

Iterative solvers and low-rank tensor approximations

Given a low-rank format $\mathcal{M}_r$, replace standard iterations

$$u_n = B_n(u_{n-1})$$

with

$$u_n \in \mathcal{M}_{r_n} \quad \text{s.t.} \quad \|u_n - B_n(u_{n-1})\| \leq \epsilon$$

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

See e.g. [Kressner-Tobler 2010, Matthies-Zander 2011, Ballani-Grasedyck 2012, Bachmayr-Dahmen 2014,Giraldi-Nouy-Legrain 2014]

Direct approximation

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

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

by the optimization of a criterium

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

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

- Natural minimization problems:

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

- Residual norms:

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

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

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

Challenging issues

- Holy grail :
  - quantify optimal model order reduction : Information-based complexity
  - find algorithms achieving (quasi-)optimality
- Find suitable structures: sparsity, low-rank...
- Well-conditioned mathematical formulations for (quasi-)optimal model reduction
- Samples-based constructions: how to sample for a given approximation format ?
- Goal-oriented model order reduction: quantity of interest $Q(u)$, rare event computation, sensitivity analysis, inverse problems, ...
- Software. Minimize modifications of existing simulation codes.

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.

- **SPARSE APPROXIMATION, BEST N-TERM APPROXIMATION**


F. Nobile, R. Tempone, and C. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.


A. Cohen, R. DeVore, and C. Schwab. Convergence rates of best n-term galerkin approximations for a class of elliptic spdes. *Foundations of Computational Mathematics*, 10(6):615–646, 2010.


A. Chkifa, A. Cohen, and C. Schwab. Breaking the curse of dimensionality in sparse polynomial approximation of parametric {PDEs}. *Journal de Mathématiques Pures et Appliquées*, (0):–, 2014.

References II


J. Beck, R. Tempone, F. Nobile, And L. Tamellini. On the optimal polynomial approximation of stochastic pdes by galerkin and collocation methods. *Mathematical Models and Methods in Applied Sciences*, 22(09):1250023, 2012.


J. Beck, F. Nobile, L. Tamellini, and R. Tempone. Convergence of quasi-optimal stochastic galerkin methods for a class of PDES with random coefficients. *Computers & Mathematics with Applications*, 67(4):732 – 751, 2014.

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


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

References III


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. *SIAM Journal on Scientific Computing*, 36(4):A1850–A1870, 2014.


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. *ESAIM: Mathematical Modelling and Numerical Analysis*, eFirst, 4 2014.

- **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, R. Lebrun, A. Nouy, and P. Rai. A least-squares method for sparse low rank approximation of multivariate functions. *ArXiv e-prints*, April 2013.


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.