Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

DEPARTMENT OF CIVIL, ENVIRONMENTAL AND GEOMATIC ENGINEERING
CHAIR OF RISK, SAFETY & UNCERTAINTY QUANTIFICATION

Sparse polynomial chaos expansions for solving high-dimensional UQ problems

Bruno Sudret

Risk, Safety & Uncertainty Quantification

1st International Conference on Uncertainty Quantification in Computational Sciences and Engineering

Chair of Risk, Safety and Uncertainty quantification

The Chair carries out research projects in the field of uncertainty quantification for engineering problems with applications in structural reliability, sensitivity analysis, model calibration and reliability-based design optimization

Research topics

- Structural reliability analysis
- **Polynomial chaos expansions** and stochastic finite element methods
- Gaussian process modelling (Kriging)
- Bayesian model calibration and stochastic inverse problems
- **Global sensitivity analysis**
- Reliability-based design optimization

http://www.rsuq.ethz.ch

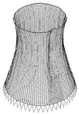
Computational models

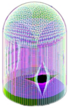
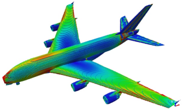
Complex engineering systems are designed using computational models that are based on:

- A mathematical description of the physics
- Numerical algorithms that solve the resulting set of (*e.g.* partial differential) equations, *e.g.* finite element models

Computational models are used:

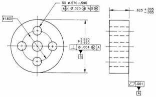
- Together with experimental data for calibration purposes
- To explore the design space ("virtual prototypes")
- To optimize the system w.r.t cost constraints
- To assess its robustness w.r.t uncertainty and its reliability



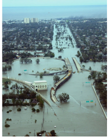



Sources of uncertainty

- Differences between the designed and the real system:
  - Dimensions (tolerances in manufacturing)
  - Material properties (*e.g.* variability of the stiffness or resistance)



- Unforecast exposures: exceptional service loads, natural hazards (earthquakes, floods), climate loads (hurricanes, snow storms, etc.)



Uncertainty quantification in engineering and applied sciences

- Uncertainty quantification arrives on top of well defined simulation procedures (legacy codes)
- State-of-the-art computational models are complex: coupled problems (thermo-mechanics), plasticity, large strains, contact, buckling, etc.
- A single simulation is already costly (*e.g.* several hours)
- The input variables modelling aleatory uncertainty are often non Gaussian. The size of the input random vector is typically 10-100

Need for non intrusive and parsimonious methods for uncertainty quantification

Outline

Polynomial chaos expansions in a nutshell

Ghanem & Spanos (1991); Xiu & Karniadakis (2002); Soize & Ghanem (2004)

- Consider the input random vector $\boldsymbol{X}$ (dim $\boldsymbol{X} = M$) with given probability density function (PDF) $f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{M} f_{X_i}(x_i)$
- Assuming that the random output $Y = \mathcal{M}(\boldsymbol{X})$ has finite variance, it can be cast as the following polynomial chaos expansion:

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

  where :

  - $y_{\boldsymbol{\alpha}}$: coefficients to be computed (coordinates)
  - $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$: basis functions

- The PCE basis $\left\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}),\, \boldsymbol{\alpha} \in \mathbb{N}^M\right\}$ is made of multivariate orthonormal polynomials

Multivariate polynomial basis

- Univariate orthogonal polynomials $\{P_k^{(i)},\ k \in \mathbb{N}\}$ are built for each input variable X_i:

$$\left\langle P_j^{(i)}(x_i), P_k^{(i)}(x_i) \right\rangle = \int P_j^{(i)}\ P_k^{(i)}\ f_{X_i}(x_i)\, dx_i = \ \gamma_j^{(i)}\ \delta_{jk}$$

- Normalization:

$$\Psi_j^{(i)} = P_j^{(i)} / \sqrt{\gamma_j^{(i)}} \quad i = 1, \ldots, M, \quad j \in \mathbb{N}$$

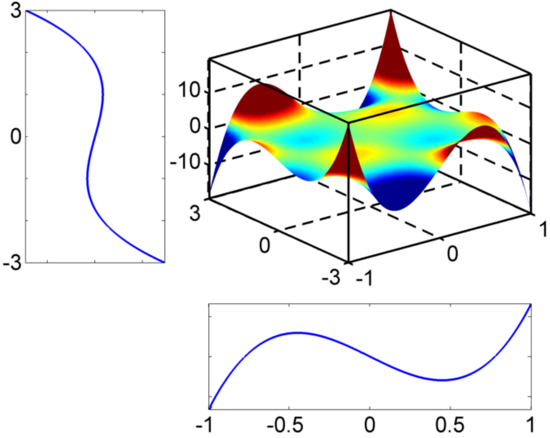
- Tensor product construction

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{i=1}^{M} \Psi_{\alpha_i}^{(i)}(x_i) \qquad \mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \Psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right] = \delta_{\boldsymbol{\alpha\beta}}$$

Example: $M = 2$

Xiu & Karniadakis (2002)

$\boldsymbol{\alpha} = [3\,,\,3]$ $\qquad \Psi_{(3,3)}(\boldsymbol{x}) = \tilde{P}_3(x_1) \cdot \tilde{He}_3(x_2)$



- $X_1 \sim \mathcal{U}(-1,1)$: Legendre polynomials
- $X_2 \sim \mathcal{N}(0,1)$: Hermite polynomials

Isoprobabilistic transform

- Classical orthogonal polynomials are defined for reduced variables, *e.g.* :
  - standard normal variables $\mathcal{N}(0, 1)$
  - standard uniform variables $\mathcal{U}(-1, 1)$
- In practical UQ problems the physical parameters are modelled by random variables that are:
  - not necessarily reduced, *e.g.* $X_1 \sim \mathcal{N}(\mu, \sigma)$, $X_2 \sim \mathcal{U}(a, b)$, etc.
  - not necessarily from a classical family, *e.g.* lognormal variable

Need for isoprobabilistic transforms

Isoprobabilistic transform

Independent variables

- Given the marginal CDFs $X_i \sim F_{X_i} \qquad i = 1, \ldots, M$
- A one-to-one mapping to reduced variables is used:

$$X_i = F_{X_i}^{-1}\left(\frac{\xi_i + 1}{2}\right) \qquad \text{if } \xi_i \sim \mathcal{U}(-1\,,\,1)$$

$$X_i = F_{X_i}^{-1}\left(\Phi(\xi_i)\right) \qquad \text{if } \xi_i \sim \mathcal{N}(0,1)$$

- The best choice is dictated by the least non linear transform

General case

Sklar's theorem (1959)

- The joint CDF is defined through its marginals and copula

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \mathcal{C}\left(F_{X_1}(x_1), \ldots, F_{X_M}(x_M)\right)$$

- Rosenblatt or Nataf isoprobabilistic transform is used

Truncation scheme

- For practical computation, a truncated series is defined:

$$Y = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

- The classical truncation scheme contains all multi-indices of total degree $|\boldsymbol{\alpha}| \stackrel{\text{def}}{=} \sum_{i=1}^{M} \alpha_i$ smaller than p

$$\mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M \ : \ |\boldsymbol{\alpha}| \leq p\} \qquad \text{card } \mathcal{A}^{M,p} \equiv P = \binom{M+p}{p}$$

$M \backslash p$	2	3	5	7	10
2	6	10	21	36	66
3	10	20	56	120	286
5	21	56	252	792	3,003
10	66				184,756
50	1,326				75,394,027,566
100	5,151	176,851	96,560,646	26,075,972,546	46,897,636,623,981

Curse of dimensionality

Outline

Various methods for computing the coefficients

Intrusive approaches

- Historical approaches: projection of the equations residuals in the Galerkin sense — Ghanem *et al.* ; Le Maître *et al.* , Babuska, Tempone *et al.* .; Karniadakis *et al.* , etc.
- Proper generalized decompositions — Nouy *et al.* , 2007-10

Non intrusive approaches

- Non intrusive methods consider the computational model $\mathcal{M}$ as a black box
- They rely upon a design of numerical experiments, *i.e.* a n-sample $\mathcal{X} = \{\boldsymbol{x}^{(i)} \in \mathcal{D}_{\boldsymbol{X}},\, i = 1, \, \dots \,, n\}$ of the input parameters
- Different classes of methods are available:
  - projection: by simulation or quadrature — Matthies & Keese, 2005; Le Maître *et al.*
  - stochastic collocation — Xiu, 2007-09; Nobile, Tempone *et al.* , 2008; Ma & Zabaras, 2009
  - least-square minimization — Berveiller *et al.* , 2006; Blatman & S., 2008-11

Statistical approach: least-square minimization

Isukapalli (1999); Berveiller *et al.* (2006)

Principle

The exact (infinite) series expansion is considered as the sum of a truncated series and a residual:

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) + \varepsilon_P \equiv \mathbf{Y}^{\mathsf{T}} \boldsymbol{\Psi}(\boldsymbol{X}) + \varepsilon_P(\boldsymbol{X})$$

where : $\mathbf{Y} = \{y_{\boldsymbol{\alpha}}, \, \boldsymbol{\alpha} \in \mathcal{A}\} \equiv \{y_0, \, \dots \, , y_{P-1}\}$ (P unknown coef.)

$$\boldsymbol{\Psi}(\boldsymbol{x}) = \{\Psi_0(\boldsymbol{x}), \, \dots \, , \Psi_{P-1}(\boldsymbol{x})\}$$

Least-square minimization

The unknown coefficients are estimated by minimizing the mean square residual error:

$$\hat{\mathbf{Y}} = \arg\min \mathbb{E}\left[\left(\mathbf{Y}^{\mathsf{T}} \boldsymbol{\Psi}(\boldsymbol{X}) - \mathcal{M}(\boldsymbol{X})\right)^2\right]$$

Least-Square Minimization: discretized solution

Ordinary least-square (OLS)

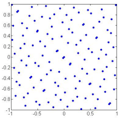
- An estimate of the mean square error (sample average) is minimized:

$$\hat{\mathbf{Y}} = \arg \min_{\mathbf{Y} \in \mathbb{R}^P} \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{Y}^\mathsf{T} \boldsymbol{\Psi}(\boldsymbol{x}^{(i)}) - \mathcal{M}(\boldsymbol{x}^{(i)}) \right)^2$$

Procedure

- Select an experimental design and evaluate the model response

$$\mathbf{M} = \left\{ \mathcal{M}(\boldsymbol{x}^{(1)}), \, \dots \, , \mathcal{M}(\boldsymbol{x}^{(n)}) \right\}^\mathsf{T}$$



- Compute the experimental matrix

$$\mathbf{A}_{ij} = \Psi_j \left(\boldsymbol{x}^{(i)} \right) \quad i = 1, \dots , n \; ; \; j = 0, \dots , P-1$$

- Solve the resulting linear system

$$\hat{\mathbf{Y}} = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\mathbf{M}$$

Simple is beautiful !

Error estimators

- In least-squares analysis, the generalization error is defined as:

$$E_{gen} = \mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{X}) - \mathcal{M}^{\mathsf{PC}}(\boldsymbol{X})\right)^2\right] \qquad \mathcal{M}^{\mathsf{PC}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

- The empirical error based on data set $\mathcal{X}$:

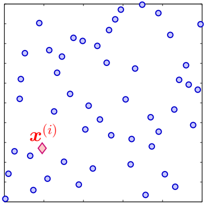
$$E_{emp} = \frac{1}{n} \sum_{i=1}^{n} \left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{\mathsf{PC}}(\boldsymbol{x}^{(i)})\right)^2$$

  is a poor estimator (overfitting):

- Model validation shall be carried out with independent data

Leave-one-out cross validation

Leave-one-out cross validation



- An experimental design $\mathcal{X} = \{\boldsymbol{x}^{(j)},\ j = 1, \ldots, n\}$ is selected
- Polynomial chaos expansions are built using all points but one, *i.e.* based on $\mathcal{X} \backslash \boldsymbol{x}^{(i)} = \{\boldsymbol{x}^{(j)},\ j = 1, \ldots, n,\ j \neq i\}$
- Leave-one-out error (PRESS)

$$E_{LOO} = \frac{1}{n} \sum_{i=1}^{n} \left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{PC \backslash i}(\boldsymbol{x}^{(i)}) \right)^2$$

- Computing directly from a single PC analysis

$$E_{LOO} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{PC}(\boldsymbol{x}^{(i)})}{1 - h_i} \right)^2$$

where h_i is the i-th diagonal term of matrix $\mathbf{A}(\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}$

Least-squares analysis: Wrap-up

Algorithm 1: OLS

Input: Computational budget n
Initialization
    Experimental design $\mathcal{X} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}\}$
    Run model $\mathcal{X} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}\}$
PCE construction
    for $p = p_{\min} : p_{\max}$ **do**
        Select candidate basis $\mathcal{A}^{M,p}$
        Solve OLS problem
        Compute $e_{\mathrm{LOO}}(p)$
    end
    $p^* = \arg\min e_{\mathrm{LOO}}(p)$
Return Best PCE of degree p^*

Post-processing sparse PC expansions

Statistical moments

- Due to the orthogonality of the basis functions $(\mathbb{E}[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\Psi_{\boldsymbol{\beta}}(\boldsymbol{X})] = \delta_{\boldsymbol{\alpha\beta}})$ and using $\mathbb{E}[\Psi_{\boldsymbol{\alpha}\neq\mathbf{0}}] = 0$ the statistical moments read:

$$\text{Mean:} \qquad \hat{\mu}_Y = y_0$$

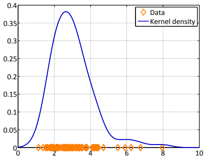
$$\text{Variance:} \qquad \hat{\sigma}_Y^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{A}\setminus\mathbf{0}} y_{\boldsymbol{\alpha}}^2$$

Distribution of the QoI

- The PCE can be used as a response surface for sampling:

$$\mathfrak{y}_j = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\,\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}_j) \quad j = 1, \ldots, n_{big}$$

- The PDF of the response is estimated by histograms or kernel smoothing



Sensitivity analysis

Goal

Sobol' (1993); Saltelli *et al.* (2000)

Global sensitivity analysis aims at quantifying which input parameter(s) (or combinations thereof) influence the most the response variability (variance decomposition)

Hoeffding-Sobol' decomposition

$(\boldsymbol{X} \sim \mathcal{U}([0,1]^M))$

$$\mathcal{M}(\boldsymbol{x}) = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(x_i) + \sum_{1 \leq i < j \leq M} \mathcal{M}_{ij}(x_i,\, x_j) + \cdots + \mathcal{M}_{12\ldots M}(\boldsymbol{x})$$

$$= \mathcal{M}_0 + \sum_{\mathbf{u} \subset \{1,\, \ldots\, ,M\}} \mathcal{M}_{\mathbf{u}}(\boldsymbol{x}_{\mathbf{u}}) \qquad (\boldsymbol{x}_{\mathbf{u}} \stackrel{\text{def}}{=} \{x_{i_1},\, \ldots\, , x_{i_s}\})$$

- The summands satisfy the orthogonality condition:

$$\int_{[0,1]^M} \mathcal{M}_{\mathbf{u}}(\boldsymbol{x}_{\mathbf{u}})\, \mathcal{M}_{\mathbf{v}}(\boldsymbol{x}_{\mathbf{v}})\, d\boldsymbol{x} = 0 \qquad \forall\, \mathbf{u} \neq \mathbf{v}$$

Sobol' indices

Total variance:

$$D \equiv \mathrm{Var}\left[\mathcal{M}(\boldsymbol{X})\right] = \sum_{\mathbf{u} \subset \{1, \dots, M\}} \mathrm{Var}\left[\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}})\right]$$

- Sobol' indices:

$$S_{\mathbf{u}} \stackrel{\text{def}}{=} \frac{\mathrm{Var}\left[\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}})\right]}{D}$$

- First-order Sobol' indices:

$$S_i = \frac{D_i}{D} \qquad D_i = \mathrm{Var}_{X_i}\left[\mathcal{M}_i(X_i)\right]$$

Quantify the additive effect of each input parameter separately

- Total Sobol' indices:

$$S_i^T \stackrel{\text{def}}{=} \sum_{\mathbf{u} \supset i} S_{\mathbf{u}}$$

Quantify the total effect of X_i, including interactions with the other variables.

Link with PC expansions

Sobol decomposition of a PC expansion

Sudret, RESS (2006-08)

Obtained by reordering the terms of the (truncated) PC expansion
$\mathcal{M}^{\text{PC}}(\boldsymbol{X}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$

Interaction sets

$$\forall\, \mathbf{u} \stackrel{\text{def}}{=} \{i_1, \ldots, i_s\} : \qquad \mathcal{A}_{\mathbf{u}} = \{\boldsymbol{\alpha} \in \mathcal{A} \,:\, k \in \mathbf{u} \Leftrightarrow \alpha_k \neq 0\}$$

$$\mathcal{M}^{\text{PC}}(\boldsymbol{x}) = \mathcal{M}_0 + \sum_{\mathbf{u}\subset\{1,\ldots,M\}} \mathcal{M}_{\mathbf{u}}(\boldsymbol{x}_{\mathbf{u}}) \qquad \text{where} \qquad \mathcal{M}_{\mathbf{u}}(\boldsymbol{x}_{\mathbf{u}}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{\alpha}\in\mathcal{A}_{\mathbf{u}}} y_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$$

PC-based Sobol' indices

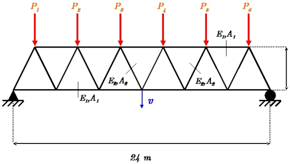
$$S_{\mathbf{u}} = D_{\mathbf{u}}/D = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_{\mathbf{u}}} y_{\boldsymbol{\alpha}}^2 \Big/ \sum_{\boldsymbol{\alpha}\in\mathcal{A}\setminus\mathbf{0}} y_{\boldsymbol{\alpha}}^2$$

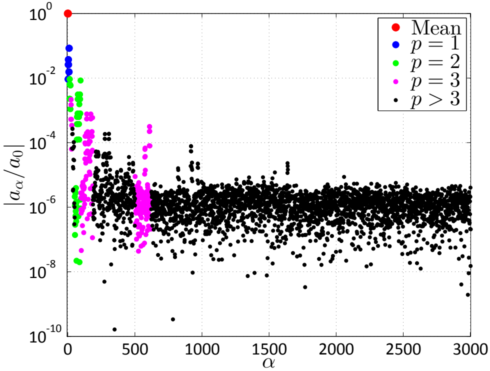
The Sobol' indices are obtained analytically, at any order from the coefficients of the PC expansion

Outline

Why are sparse representations relevant?

- Elastic truss structure
- $M = 10$ independent input variables (loads / Young's moduli / cross sections)
- PCE of degree $p = 5$ (3,003 coefficients)





Low-rank truncation schemes

Ockham's razor

"entia non sunt multiplicanda praeter necessitatem" (entities must not be multiplied beyond necessity) W. Ockham (c. 1287-1347)

Sparsity-of-effects principle

In most engineering problems, only low-order interactions between the input variables are relevant.

Use of low-rank monomials

Definition

The rank of a multi-index $\boldsymbol{\alpha}$ is the number of active variables of $\Psi_{\boldsymbol{\alpha}}$, *i.e.* the number of non-zero terms in $\boldsymbol{\alpha}$:

$$||\boldsymbol{\alpha}||_0 = \sum_{i=1}^{M} \mathbf{1}_{\{\alpha_i > 0\}}$$

Hyperbolic truncation sets

Definition

Blatman (2009); Blatman & Sudret, J. Comp. Phys (2011)

- The $q-$norm of a multi-index $\boldsymbol{\alpha}$ is defined by:

$$||\boldsymbol{\alpha}||_q \equiv \left(\sum_{i=1}^{M} \alpha_i^q \right)^{1/q} \quad , \quad 0 < q \leq 1$$

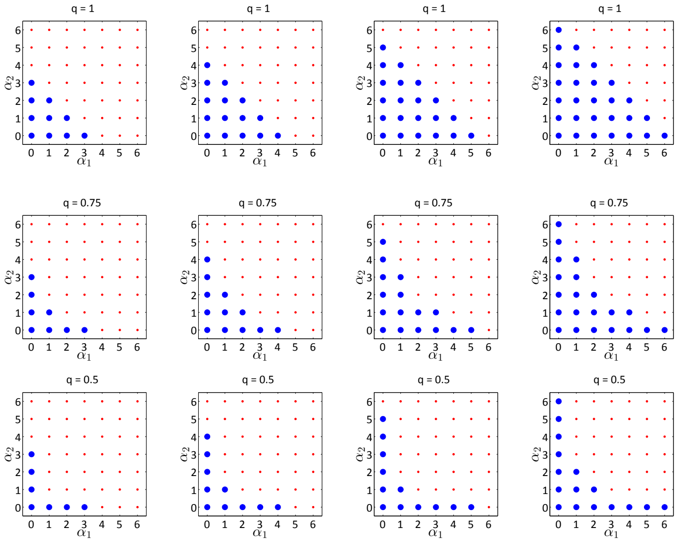
- The hyperbolic truncation sets read:

$$\mathcal{A}_q^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M \ : \ ||\boldsymbol{\alpha}||_q \leq p\}$$

Limit cases

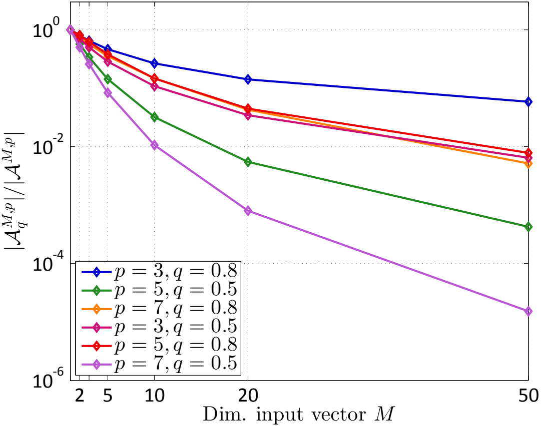
- $q = 1$: standard truncation scheme (all polynomials of maximal total degree p)
- $q \to 0$: additive model (no interaction)

Hyperbolic truncation sets



Hyperbolic truncation sets

- For a given value of $0 < q \leq 1$, the index of sparsity tends to zero when M and p increase



$$IS = \frac{\left|\mathcal{A}_q^{M,p}\right|}{\left|\mathcal{A}^{M,p}\right|}$$

How to get sparse expansions?

Blatman & Sudret, JCP (2011)

- Sparsity in the solution can be induced by ℓ_1-regularization:

$$\boldsymbol{y}_{\boldsymbol{\alpha}} = \arg\min \frac{1}{n} \sum_{i=1}^{n} \left(\mathsf{Y}^{\mathsf{T}} \boldsymbol{\Psi}(\boldsymbol{x}^{(i)}) - \mathcal{M}(\boldsymbol{x}^{(i)}) \right)^2 + \lambda \parallel \boldsymbol{y}_{\boldsymbol{\alpha}} \parallel_1$$

- Different algorithms: LASSO, (Bayesian) compressive sensing

Doostan & Owhadi (2011); Ian, Guo, Xiu (2012); Sargsyan *et al.* (2014); Jakeman, Eldred, Sargsyan (2015)

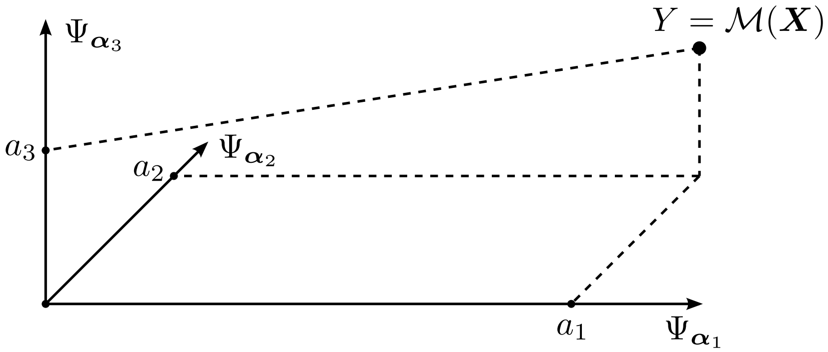
Least Angle Regression

Efron *et al.* (2004)

- Least Angle Regression (LAR) solves the LASSO problem for different values of the penalty constant in a single run
- The various PC expansions obtained have $1, 2, \ldots, \min(n, |\mathcal{A}|)$ terms

Least angle regression
Implementation

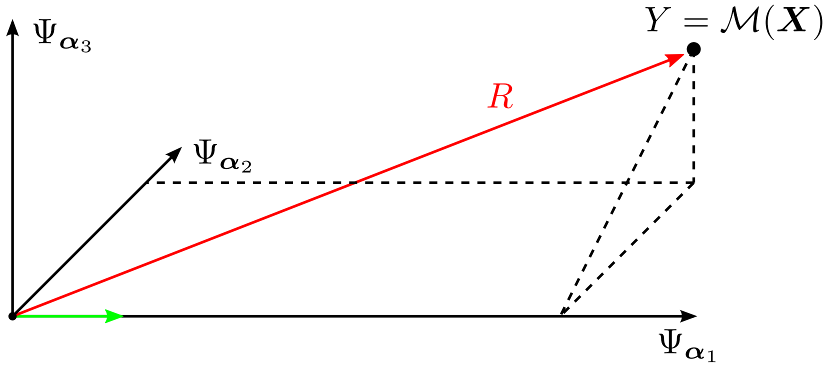
Efron *et al.* , 2004



Consider a 3-dimensional vector

Least angle regression
Implementation

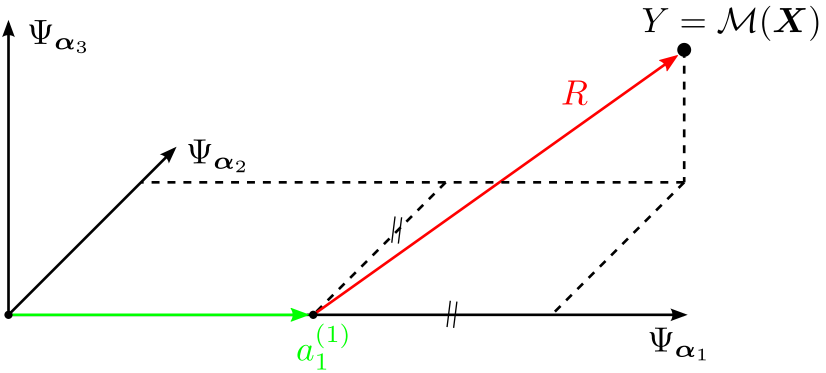
Efron *et al.* , 2004



- The algorithm is initialized with $Y^{(0)} = 0$. The residual is $R = Y = \mathcal{M}(\boldsymbol{X})$
- The most correlated regressor is $\Psi_{\boldsymbol{\alpha}_1}$

Least angle regression
Implementation

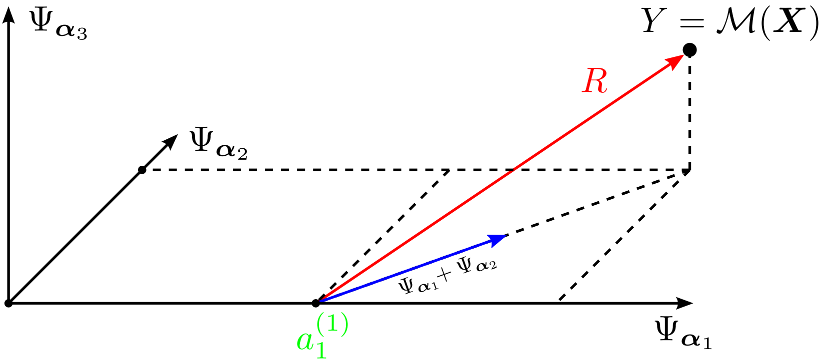
Efron *et al.* , 2004



- A move in the direction $\Psi_{\boldsymbol{\alpha}_1}$ is carried out so that the residual $Y - a_1^{(1)}\,\Psi_{\boldsymbol{\alpha}_1}$ becomes equicorrelated with $\Psi_{\boldsymbol{\alpha}_1}$ and $\Psi_{\boldsymbol{\alpha}_2}$
- The 1-term sparse approximation of Y is $a_1^{(1)}\,\Psi_{\boldsymbol{\alpha}_1}$

Least angle regression
Implementation

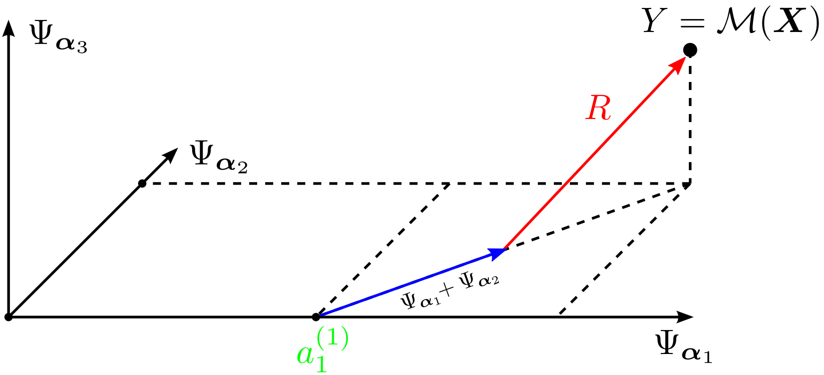
Efron *et al.* , 2004



- A move is jointly made in the direction $\Psi_{\boldsymbol{\alpha}_1} + \Psi_{\boldsymbol{\alpha}_2}$ until the residual becomes equicorrelated with a third regressor

Least angle regression
Implementation

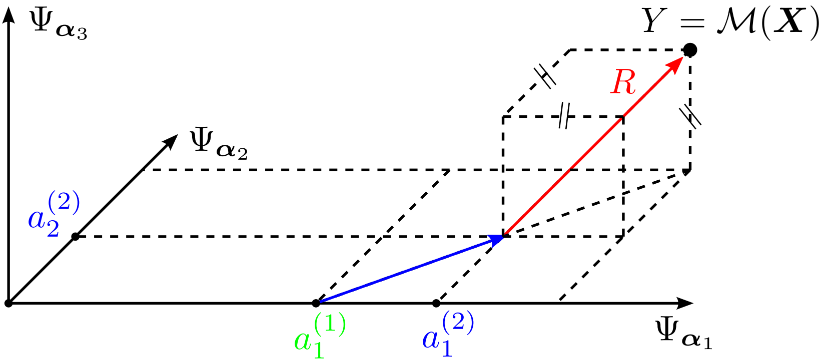
Efron *et al.* , 2004



- A move is jointly made in the direction $\Psi_{\boldsymbol{\alpha}_1} + \Psi_{\boldsymbol{\alpha}_2}$ until the residual becomes equicorrelated with a third regressor

Least angle regression
Implementation

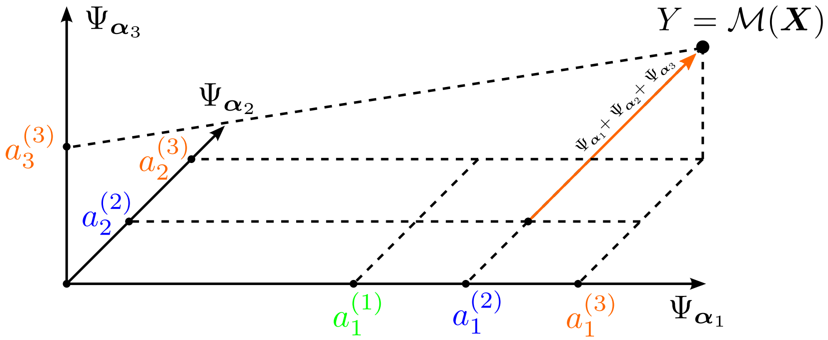
Efron *et al.* , 2004



- This gives the 2-term sparse approximation

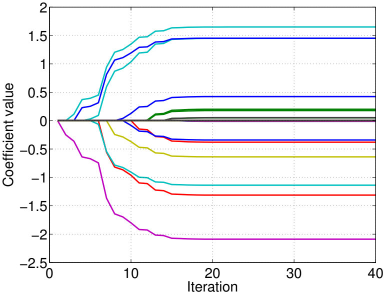
Least angle regression
Implementation

Efron *et al.* , 2004



- etc.
- In finite dimension, LAR eventually yields the same results as projection in P steps

Least angle regression
Path of solutions



- A path of solutions is obtained containing $1, 2, .., \min(n, |\mathcal{A}|)$ terms.
- Leave-one-out error E_{LOO} is computed for each solution and the best model (smallest error) is selected

$$E_{LOO} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{PC}(\boldsymbol{x}^{(i)})}{1 - h_i} \right)^2$$

h_i: i-th diagonal term of matrix $\mathbf{A}(\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}$ and $\mathbf{A}_{ij} = \Psi_j(\boldsymbol{x}^{(i)})$

Sparse PCE: wrap up

Algorithm 2: LAR-based Polynomial chaos expansion

Input: Computational budget n

Initialization

  Sample experimental design $\mathcal{X} = \{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)}\}$

  Evaluate model response $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{x}^{(1)}), \dots, \mathcal{M}(\boldsymbol{x}^{(n)})\}$

PCE construction

  for $p = p_{\min} : p_{\max}$ **do**

    for $q \in \mathcal{Q}$ **do**

      Select candidate basis $\mathcal{A}_q^{M,p}$

      Run LAR for extracting the optimal sparse basis $\mathcal{A}^*(p,q)$

      Compute coefficients $\{y_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathcal{A}^*(p,q)\}$ by OLS

      Compute $e_{\mathsf{LOO}}(p,q)$

    end

  end

  $(p^*, q^*) = \arg\min e_{\mathsf{LOO}}(p,q)$

Return Optimal sparse basis $\mathcal{A}^*(p,q)$, PCE coefficients, $e_{\mathsf{LOO}}(p^*,q^*)$

Tolerance-driven sparse PCE: wrap up

Algorithm 3: Tolerance-driven Sparse PCE

Input
 Initial and max. computational budget n_{ini}, n_{max} batch size B
 Target error TOL
Initialization
 Apply LARbasedPCE(n_{ini}), return $e_{\mathrm{LOO}}(n_{ini})$
Enrich ED
 $n \leftarrow n_{ini}$
 while ($e_{LOO}(n) > TOL$) **&** $(n + B \leq n_{max})$ **do**
  Enrich ED: $\mathcal{X} \leftarrow \mathcal{X} \cup \{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(B)}\}$
  $n \leftarrow n + B$
  Apply LARbasedPCE(n)
 end
Return Final ED size n, optimal sparse basis and PCE coefficients, $e_{\mathrm{LOO}}(n)$

Outline

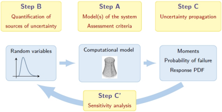
The UQLab framework

UQLab ...

... The Uncertainty Quantification Laboratory

"Make uncertainty quantification available for anybody, in any field of applied science and engineering"

- Matlab-based core managing system (MODEL / INPUT / ANALYSIS objects)

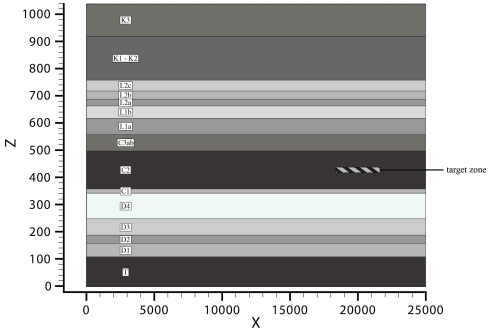


- Modules: surrogate models (Gaussian processes / polynomial chaos expansions), sensitivity analysis, reliability analysis
- Dispatcher to HPC infrastructure

Geological model

Joint work with University of Neuchâtel

Deman, Konakli, BS, Kerrou, Perrochet & Benabderrahmane, Reliab. Eng. Sys. Safety (submitted)



- Idealized model of the Paris Basin
- Two-dimensional cross section

  (25 km long / 1,040 m depth) with 5×5 m mesh (10^6 elements)
- 15 homogeneous layers

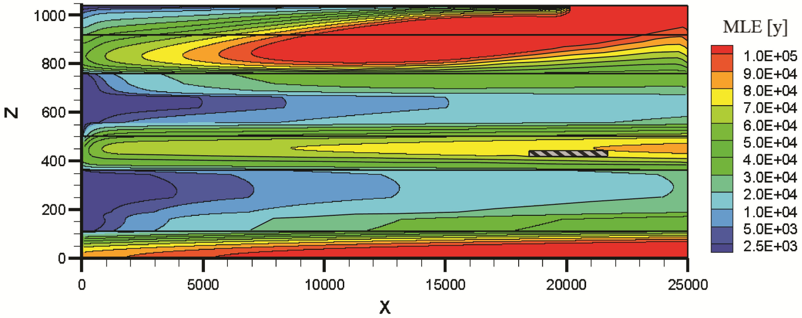
- Steady-state flow with Dirichlet boundary conditions:

$$\nabla \cdot (\mathbf{K} \cdot \nabla H) = 0$$

Mean life-time expectancy

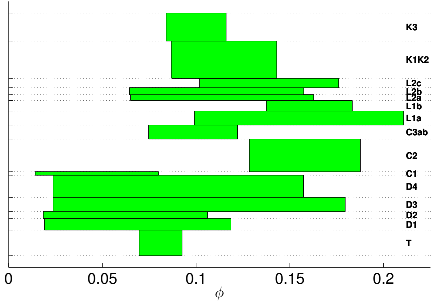
Definition

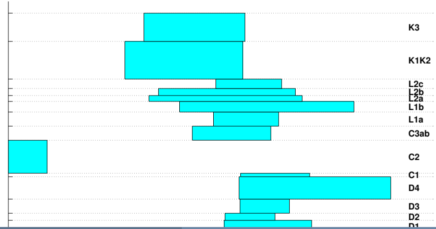
The Mean Lifetime Expectancy $MLE(x)$ is the time required for a molecule of water at point x to get out of the boundaries of the model



Map of mean lifetime expectancy (nominal case)

Probabilistic model of porosity / conductivity





Nominal conductivity (K_x) vs. porosity

Layer	K_x [m/s]	ϕ [-]
K3	9.01E-09	0.0100
K1-K2	4.53E-09	0.1150
L2c	1.10E-06	0.1389
L2b	3.46E-07	0.1110
L2a	1.62E-07	0.1139
L1b	1.49E-05	0.1604
L1a	1.17E-06	0.1549
C3ab	4.59E-08	0.0984
C2	1.99E-13	0.1580
C1	1.89E-06	0.0470
D4	1.65E-05	0.0905
D3	1.76E-06	0.1016
D2	2.62E-07	0.0623
D1	3.23E-06	0.0688
T	1.95E-12	0.0810

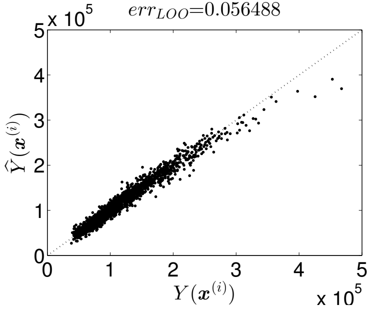
Other parameters

Parameter	Notation	Range
Porosity	$\phi^i,\ i = 1, \ldots, 15$	$[\phi^i_{min}\,,\ \phi^i_{max}]$
Anisotropy of hydraulic conductivity tensor	$A^i_K,\ i = 1, \ldots, 15$	$[0.01\,,\ 1]$
Euler angle of hydraulic conductivity tensor	$\theta^i,\ i = 1, \ldots, 15$	$[-30\,,\ 30](^\circ)$
Longitudinal component of dispersivity tensor	$\alpha^i_L,\ i = 1, \ldots, 15$	$[5\,,\ 25]$
Anisotropy of dispersivity tensor	$A^i_\alpha,\ i = 1, \ldots, 15$	$[5\,,\ 25]$
Hydraulic gradient $(10^{-3} m/m)$		
Dogger sequence	∇H_D	[0.64 , 0.96]
Oxfordian sequence	∇H_O	[2.40 , 3.60]
Top of the model	∇H_{top}	[2.72 , 4.08]

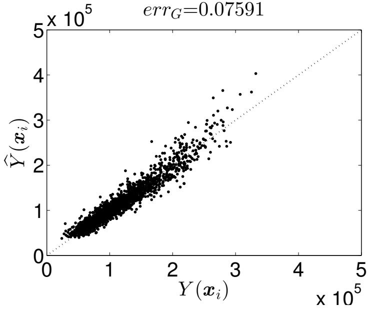
78 independent variables with uniform distributions

Polynomial chaos expansions

- Experimental design of size 2,000 (Maximin Latin Hypercube Sampling). Independent validation set of size 2,000
- Truncation scheme: $p = 8$, $q = 0.5$
- Sparse basis size: 185 / Full-basis size 5.3×10^{10}. Only 68 out of 78 parameters are included

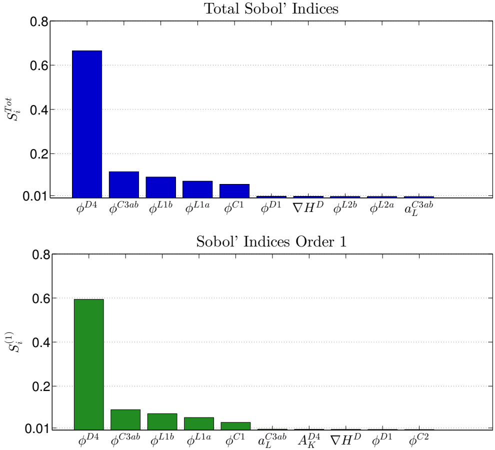


Training set



Validation set

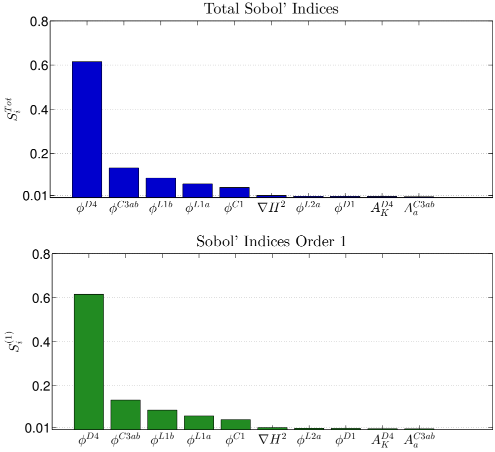
Sobol' sensitivity indices



Parameter	$\sum_j S_j^{(1)}$
ϕ	0.8664
A_K	0.0088
θ	0.0029
α_L	0.0076
A_α	0.0000
∇H	0.0057

- Uncertainties on the porosities (and associated conductivities) drive the MLE uncertainty
- Second-order effects have been identified

Sobol' sensitivity indices: using 200 model runs



Only 200 model runs allow one to detect the important parameters out of 78

Outline

Problem statement

Premise: In case of time-dependent governing equations, the response of the system is a time-dependent function:

$$Y(t) = \mathcal{M}(\boldsymbol{X};\, t)$$

- ordinary differential equations with random coefficients (chemical reactions)
- fluid dynamics
- structural dynamics (*e.g.* earthquake engineering)

Time-frozen PCE

- Consider the discretized deterministic solutions n_{TS} time steps:

$$y_i(t_j) = \mathcal{M}(\boldsymbol{x}^{(i)};\, t_j) \qquad i = 1, \ldots, n, \; j = 1, \ldots, n_{TS}$$

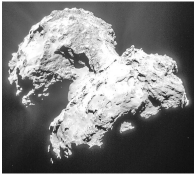
- Build up PCE independently at each time-instant (considered frozen)

$$Y(t_j) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{t_j}} y_{\boldsymbol{\alpha}}(t_j)\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

Fails due to increasing complexity of the input/output map when $t \rightarrow +\infty$

Example: rigid body dynamics

Mai & S., MascotNum Workshop, 2015



http://www.esa.int/spaceinimages/Missions/Rosetta

Rotation of a rigid body described by Euler's equations

$$\begin{cases} M_x = I_{xx}\,\dot{\omega}_x - (I_{yy} - I_{zz})\,\omega_y\,\omega_z \\ M_y = I_{yy}\,\dot{\omega}_y - (I_{zz} - I_{xx})\,\omega_z\,\omega_x \\ M_z = I_{zz}\,\dot{\omega}_z - (I_{xx} - I_{yy})\,\omega_x\,\omega_y \end{cases}$$

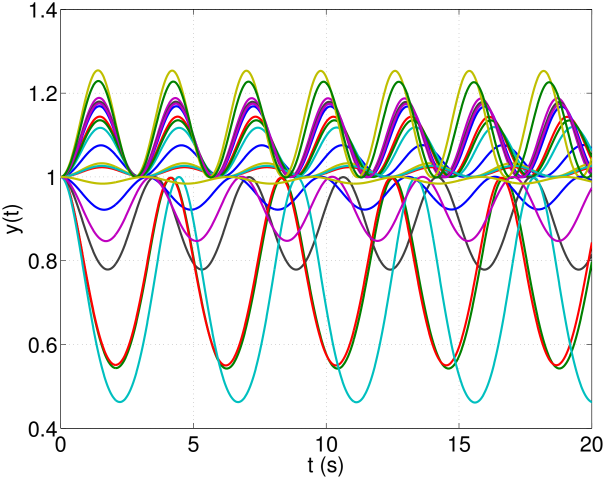
Reduced system

- $M_x = M_y = M_z = 0$
- $x(0) = 0,\ y(0) = 1,\ z(0) = 1$
- $I_{xx} = \frac{1-c}{2} I_{yy},\ I_{zz} = \frac{1+c}{2} I_{yy}$

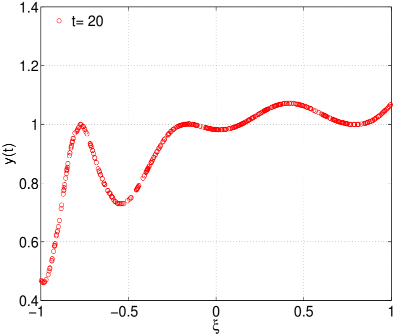
$$\begin{cases} \dot{x} = yz \\ \dot{y} = c \times xz \\ \dot{z} = -xy \end{cases}$$

where $c \sim \mathcal{U}(-0.8\,,\,0.6)$

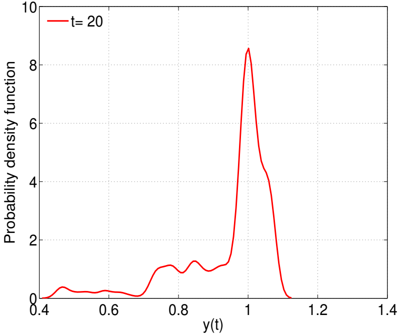
Different trajectories for various values of c



Stochastic dependence $y(c,\ t)$ for different time instants t

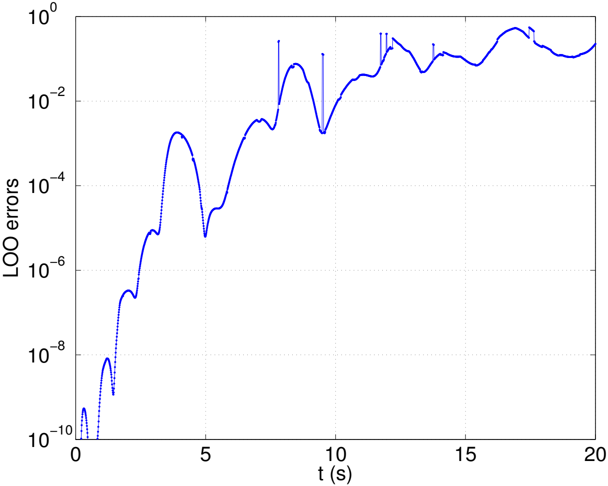


$y(c,\ t = 20)$



PDF of $Y_{t=20}$

Time-frozen PCE: LOO error



Stochastic time warping

Heuristics

Le Maître *et al.* (2010)

Introduce a virtual time scale τ such that the current trajectory $y(\boldsymbol{x}^{(i)}, \tau)$ is "similar" to a reference trajectory

Measure of dissimilarity

$$\text{diss}\left[y(t)\,,\,y_{ref}(t)\right] \stackrel{\text{def}}{=} \frac{\left|\int_0^T y(t)\,y_{ref}(t)\,dt\right|}{\sqrt{\int_0^T y^2(t)\,dt \cdot \int_0^T y_{ref}^2(t)\,dt}}$$

- It is the cross-correlation of the two signals
- Bounded between 0 and 1

Stochastic time warping: procedure

Mai & Sudret (2015)

- Choose a reference trajectory $y_{ref}(t) = \mathcal{M}(\boldsymbol{x}_{ref},\, t)$ where *e.g.* $\boldsymbol{x}_{ref} = \mu_{\boldsymbol{X}}$
- Define a stochastic time transform:

$$\tau(\boldsymbol{X}) = k(\boldsymbol{X})\, t + \phi(\boldsymbol{X})$$

- For each sample trajectory $\{y_i(t),\, i = 1,\, \dots\, ,n\}$, compute the appropriate rescaling:

$$(k_i,\, \phi_i) = \arg\min_{k,\phi} \mathsf{diss}\left[\, y_i(k\, t + \phi),\, y_{ref}(t)\right]$$

- Compute a sparse PCE of the parameters of the time transform:

$$k(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} k_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \qquad \phi(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} \phi_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

Stochastic time warping: procedure

- In the virtual time scale, trajectories show much higher coherency. τ-frozen PCE expansions apply:

$$y(\boldsymbol{X}, \tau) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}}(\tau)\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

Predictions for a new sample $\boldsymbol{x}^{(0)}$

- Predict the trajectory in the virtual time scale

$$y(\boldsymbol{x}^{(0)}, \tau) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}}(\tau)\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}^{(0)})$$

- Predict the proper time warping:

$$\tau(\boldsymbol{x}^{(0)}) = k(\boldsymbol{x}^{(0)})\, t + \phi(\boldsymbol{x}^{(0)})$$

- Map back the predicted trajectory in the real time scale:

$$y(\boldsymbol{x}^{(0)}, t) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \left(k(\boldsymbol{x}^{(0)})\, t + \phi(\boldsymbol{x}^{(0)}) \right)\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}^{(0)})$$

Application – non linear Duffing oscillator

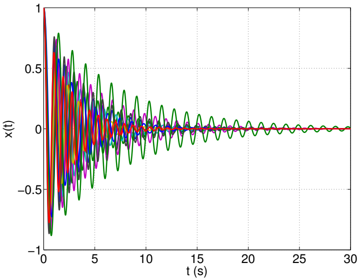
Non-linear SDOF Duffing oscillator:

$$\ddot{x}(t) + 2\,\omega\,\zeta\,\dot{x}(t) + \omega^2\,\left(x(t) + \epsilon\,x^3(t)\right) = 0$$

Initial conditions: $x(0) = 1, \quad \dot{x}(0) = 0$

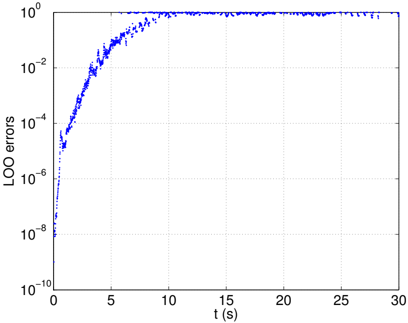
Input: 3 uniform random variables

$$\begin{aligned}
\zeta &= 0.05(1 + 0.05\,\xi_1), & \xi_1 &\sim \mathcal{U}(-1,1)\\
\omega &= 2\pi(1 + 0.2\,\xi_2), & \xi_2 &\sim \mathcal{U}(-1,1)\\
\epsilon &= -0.5(1 + 0.5\,\xi_3), & \xi_3 &\sim \mathcal{U}(-1,1)
\end{aligned}$$

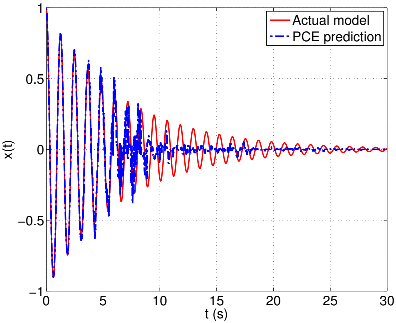


Samples of trajectories

Time-frozen PCE

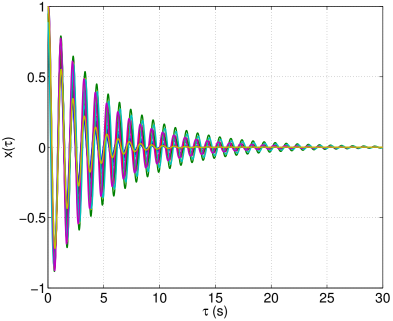


LOO error vs. time

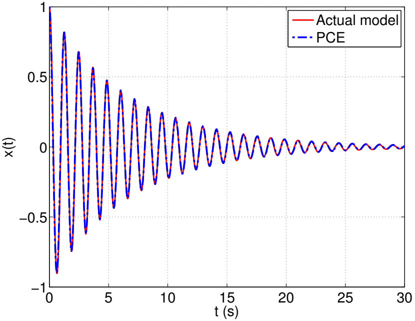


Predicted trajectory

Time-warped PCE



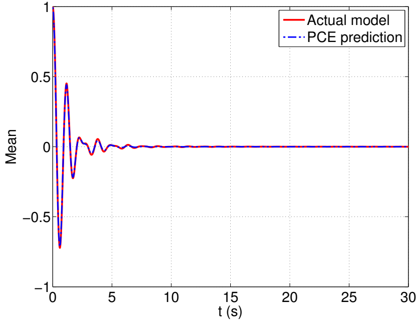
Rescaled trajectories



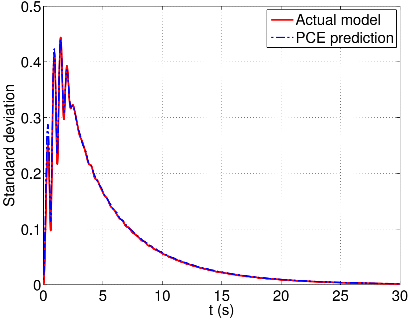
Predicted trajectory in real time t

Validation: mean and standard deviation (time-warping PCE)

Validation set: 10,000 Monte Carlo samples



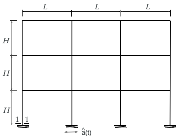
Mean response

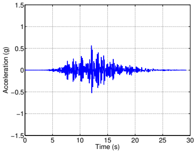


Standard deviation

Earthquake engineering applications

Mai & Sudret (2015)





- Structural systems under earthquake excitation
- Parametrized input signal in high dimension:

$$\hat{a}(t) = \alpha_1 t^{\alpha_2 - 1} \exp(-\alpha_3 t) \sum_{i=1}^{n} s_i\left(t, \boldsymbol{\lambda}(t_i)\right) U_i$$

Goal

Predict the output trajectories through time-variant PCE, *e.g.* the interstorey drift

Conclusions

- Polynomial chaos expansions are a versatile tool for solving engineering uncertainty quantification problems
- Sparse expansions are extremely efficient for global sensitivity analysis (*e.g.* $\sim$ x00 model runs for 50-100 input variables)
- An a posteriori built-in error estimator is available through leave-one-out cross validation, leading to adaptive methods (incl. adaptive experimental designs)
- Ingredients such as isoprobabilistic transforms, least-square analysis and low-rank truncation schemes are easy to understand
- … and easy to implement in a general-purpose software

Outlook and ongoing projects

- More compact representations: low-rank tensor approximations

  Chevreuil *et al.* (2013), Konakli & Sudret, UNCECOMP'2015

- Optimal small size experimental designs and local error estimation: Polynomial-chaos based Kriging

  Kersaudy *et al.* , JCP (2015) ; Schöbi & Sudret, IJUQ (2015)

- PCE expansions in case of imprecise probability description of the input parameters through p-boxes

  Schöbi & Sudret, ICASP (2015)

- Spectral likelihood expansions for solving Bayesian inverse problems

  Nagel & Sudret, PANACM (2015)

Questions ?

Acknowledgements:
K. Konakli, C.V. Mai, C. Lataniotis, S. Marelli, R. Schöbi

Thank you very much for your attention !

Chair of Risk, Safety & Uncertainty Quantification

`http://www.rsuq.ethz.ch`