Uncertainty propagation using polynomial chaos expansions

Bruno Sudret

Chair of Risk, Safety and Uncertainty Quantification

ETH Zurich

Weimar, September 5th, 2016

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

- Uncertainty modelling for engineering systems
- Structural reliability analysis
- Surrogate models (polynomial chaos expansions, Kriging, support vector machines)
- Bayesian model calibration and stochastic inverse problems
- Global sensitivity analysis
- Reliability-based design optimization

http://www.rsuq.ethz.ch

Computational models in engineering

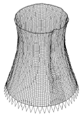
Complex engineering systems are designed and assessed using computational models, a.k.a simulators

A computational model combines:

- A mathematical description of the physical phenomena (governing equations), *e.g.* mechanics, electromagnetism, fluid dynamics, etc.

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \end{aligned}$$

- Discretization techniques which transform continuous equations into linear algebra problems

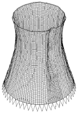


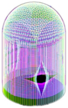
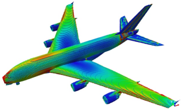
- Algorithms to solve the discretized equations

Computational models in engineering

Computational models are used:

- Together with experimental data for calibration purposes
- To explore the design space ("virtual prototypes")
- To optimize the system (*e.g.* minimize the mass) under performance constraints
- To assess its robustness w.r.t uncertainty and its reliability



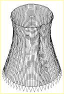



Computational models: the abstract viewpoint

A computational model may be seen as a black box program that computes quantities of interest (QoI) (a.k.a. model responses) as a function of input parameters

Vector of input parameters $\boldsymbol{x} \in \mathbb{R}^M$ — Computational model $\mathcal{M}$ — Model response $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^Q$

- Geometry
- Material properties
- Loading

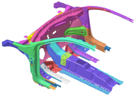


- Analytical formula
- Finite element model
- Comput. workflow

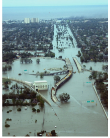
- Displacements
- Strains, stresses
- Temperature, etc.

Real world is uncertain

- 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, landslides), climate loads (hurricanes, snow storms, etc.), accidental human actions (explosions, fire, etc.)



Outline

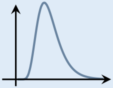
Global framework for uncertainty quantification

Step B
Quantification of sources of uncertainty

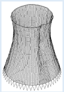
Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables



Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

B. Sudret, *Uncertainty propagation and sensitivity analysis in mechanical models – contributions to structural reliability and stochastic spectral methods (2007)*

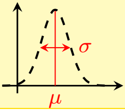
Step C: uncertainty propagation

Goal: estimate the uncertainty / variability of the quantities of interest (QoI) $Y = \mathcal{M}(\boldsymbol{X})$ due to the input uncertainty $f_{\boldsymbol{X}}$

- Output statistics, *i.e.* mean, standard deviation, etc.

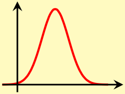
$$\mu_Y = \mathbb{E}_{\boldsymbol{X}}\left[\mathcal{M}(\boldsymbol{X})\right]$$
$$\sigma_Y^2 = \mathbb{E}_{\boldsymbol{X}}\left[\left(\mathcal{M}(\boldsymbol{X}) - \mu_Y\right)^2\right]$$

Mean/std. deviation



- Distribution of the QoI

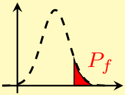
Response PDF



- Probability of exceeding an admissible threshold y_{adm}

$$P_f = \mathbb{P}\left(Y \geq y_{adm}\right)$$

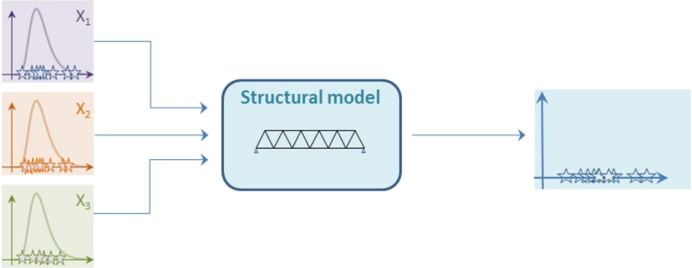
Probability of failure



Uncertainty propagation using Monte Carlo simulation

Principle: Generate virtual prototypes of the system using random numbers

- A sample set $\mathcal{X} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ is drawn according to the input distribution $f_{\boldsymbol{X}}$
- For each sample the quantity of interest (resp. performance criterion) is evaluated, say $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{x}_1), \dots, \mathcal{M}(\boldsymbol{x}_n)\}$



- The set of quantities of interest is used for moments-, distribution- or reliability analysis

Advantages/Drawbacks of Monte Carlo simulation

Advantages

- Universal method: only rely upon sampling random numbers and running repeatedly the computational model
- Sound statistical foundations: convergence when $N_{MCS} \to \infty$
- Suited to High Performance Computing: "embarrassingly parallel"

Drawbacks

- Statistical uncertainty: results are not exactly reproducible when a new analysis is carried out (handled by computing confidence intervals)
- Low efficiency: convergence rate $\propto 1/\sqrt{N_{MCS}}$

The "scattering" of Y is investigated point-by-point: if two samples $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ are almost equal, two independent runs of the model are carried out

Spectral approach

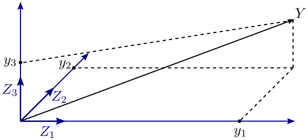
Heuristic

Instead of considering the random output $Y = \mathcal{M}(\boldsymbol{X})$ through samples, *i.e.* $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{x}_i),\ i = 1, \ldots, n\}$, Y is represented by a series expansion

$$Y = \sum_{j=0}^{+\infty} y_j \, Z_j$$

where:

- $\{Z_j\}_{j=0}^{+\infty}$ is a numerable set of random variables that forms a basis of a suitable space $\mathcal{H} \supset Y$
- $\{y_j\}_{j=0}^{+\infty}$ is the set of coordinates of Y in this basis



Spectral approach

Questions to solve

- What is the relevant mathematical framework (*i.e.* abstract space $\mathcal{H}$) to represent random variables $Y = \mathcal{M}(\boldsymbol{X})$?
- How to construct a basis of this space of $\{Z_j\}_{j=0}^{+\infty}$?
- How to compute the coefficients ? (truncation scheme)
- How to intepret the results in terms of meaningful engineering quantities ?

Outline

Polynomial chaos expansions in a nutshell

Ghanem & Spanos (1991); Sudret & Der Kiureghian (2000); 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 :

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

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

Orthogonal polynomials

Definition

- A monic polynomial of degree n reads:

$$p_n(x) = x^n + a_{n-1}\, x^{n-1} + \cdots + a_0$$

- A sequence of monic polynomials $\{\pi_k,\, k \geq 0\}$ is orthogonal with respect to a weight function $w : x \in \mathcal{D}_X \mapsto \mathbb{R}^+$ if:

$$< \pi_k\,,\, \pi_l >_w \stackrel{\text{def}}{=} \int_{\mathcal{D}_X} \pi_k(x)\, \pi_l(x)\, w(x)\, dx = \gamma_k^2\, \delta_{kl}$$

δ_{kl}: Kronecker symbol

that is:

$$< \pi_k\,,\, \pi_l >_w = 0 \quad \text{if} \quad k \neq l$$
$$< \pi_k\,,\, \pi_k >_w = \| \pi_k \|_w^2 = \gamma_k^2$$

Orthogonal polynomials

Canonical representation

- The sequence of powers $\left\{1,\, x,\, x^2, \ldots\right\}$ forms a basis of the space of polynomials.
- This basis is however not orthogonal with respect to classical weight functions

Example

Consider a uniform random variable $\mathcal{U}(-1, 1)$ with PDF $w(x) = 1/2,\, x \in [-1, 1]$ and 0 otherwise:

$$< x^p \,,\, x^q >_w = \int_{-1}^{1} x^{p+q} \, \frac{dx}{2} = \frac{1}{p+q+1} \qquad \text{if } p+q \text{ even}$$

The set of powers does NOT form an orthogonal basis:

$$< x^p \,,\, x^q >_w \neq 0$$

Basis of orthogonal polynomials

- Given the weight function w, there is a unique infinite sequence of monic orthogonal polynomials $\{\pi_k,\ k \geq 0\}$ where $\pi_0(x) \stackrel{\text{def}}{=} 1$
- This sequence may be built by the Gram-Schmidt orthogonalization procedure
- It satisfies a 3-term recurrence relation:

$$\pi_{k+1}(x) = (x - \alpha_k)\, \pi_k(x) - \beta_k\, \pi_{k-1}(x)$$

where:

$$\alpha_k = \frac{< x\, \pi_k\, ,\ \pi_k >_w}{< \pi_k\, ,\ \pi_k >_w}$$

$$\beta_k = \frac{< \pi_k\, ,\ \pi_k >_w}{< \pi_{k-1}\, ,\ \pi_{k-1} >_w}$$

Classical orthogonal polynomials

- Classical families of orthogonal polynomials have been discovered historically when solving various problems of physics, quantum mechanics, etc.
- The name of the researcher who first investigated their properties is attached to them.

$\mathcal{D}_X$	Distribution	PDF $f_X \equiv w$	Family
$[-1, 1]$	Uniform	1/2	Legendre
$\mathbb{R}$	Gaussian	$e^{-x^2/2}/\sqrt{2\pi}$	Hermite
$\mathbb{R}^+$	Exponential	e^{-x}	Laguerre
$[-1, 1]$	Beta	$\dfrac{(1-x)^\alpha(1+x)^\beta}{B(\alpha+1, \beta+1)}$	Jacobi


A.-M. Legendre (1752-1833)


C. Hermite (1822-1901)

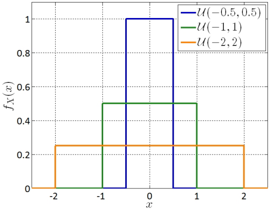

E. Laguerre (1834-1886)


C. Jacobi (1804-1851)

Legendre polynomials

Legendre polynomials are defined over $[-1, 1]$ so as to be orthogonal with respect to the uniform distribution:

$$w(x) = 1/2 \qquad x \in [-1, 1]$$



- Notation: $P_n(x),\ n \in \mathbb{N}$
- 3-term recurrence

$$P_0(x) = 1 \quad ; \quad P_1(x) = x$$

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

- P_n is solution of the ordinary differential equation

$$\left[(1 - x^2)\, P_n'(x)\right]' + n(n+1)\, P_n(x) = 0$$

First Legendre polynomials

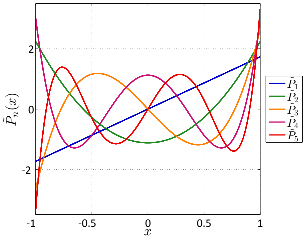
- The norm of the n-th Legendre polynomial reads:

$$\| P_n \|^2 = < P_n \, , \, P_n > = \int_{-1}^{1} P_n^2(x) \cdot \frac{1}{2} \mathrm{d}x = \frac{1}{2n+1}$$

- The orthonormal Legendre polynomials read:

$$\tilde{P}_n(x) = \sqrt{2n+1} \, P_n(x)$$

n	$P_n(x)$	$\| P_n \|^2$	$\tilde{P}_n(x)$
0	1	1	1
1	x	$1/3$	$\sqrt{3}\, P_1$
2	$\frac{1}{2}(3\, x^2 - 1)$	$1/5$	$\sqrt{5}\, P_2$
3	$\frac{1}{2}(5\, x^3 - 3\, x)$	$1/7$	$\sqrt{7}\, P_3$
4	$\frac{1}{8}(35\, x^4 - 30\, x^2 + 3)$	$1/9$	$\sqrt{9}\, P_4$
5	$\frac{1}{8}(63\, x^5 - 70\, x^3 + 15\, x)$	$1/11$	$\sqrt{11}\, P_5$



Hermite polynomials

Hermite polynomials are defined over $\mathbb{R}$ so as to be orthogonal with respect to the Gaussian distribution:

$$w(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad x \in \mathbb{R}$$

- Notation: $He_n(x), \quad n \in \mathbb{N}$
- 3-term recurrence:

$$He_0(x) = 1 \quad ; \quad He_1(x) = x$$

$$He_{n+1}(x) = x\, He_n(x) - n\, He_{n-1}(x)$$

- Normalization

$$\| He_n \|^2 = \int_{-\infty}^{+\infty} He_n^2(x)\, \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = n! \qquad n! = 1 \cdot 2 \cdot 3 \dots n$$

- The orthonormal Hermite polynomials read:

$$\tilde{He}_n(x) = He_n(x)/\sqrt{n!}$$

Hermite polynomials

- He_n is solution of the ordinary differential equation:

$$He_n^{''}(x) - x\, He_n^{'}(x) + n\, He_n(x) = 0$$

and satisfies:

$$He_n(x) = (-1)^n\, e^{x^2/2} \frac{d^n}{dx^n} \left(e^{-x^2/2} \right)$$

$$He_n^{'}(x) = n\, He_{n-1}(x)$$

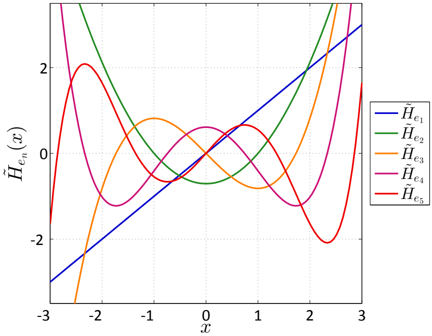
Important remark

In the literature, two families of Hermite polynomials (HP) are known:

- The "physicists' " HP are orthogonal w.r.t e^{-x^2}
- The "probabilists' " HP are orthogonal w.r.t the standard normal PDF $e^{-x^2/2}/\sqrt{2\pi}$

First Hermite polynomials

n	$He_n(x)$	$\parallel He_n \parallel^2$	$\tilde{H}e_n(x)$
0	1	1	He_0
1	x	1	He_1
2	$x^2 - 1$	2	$He_2/\sqrt{2}$
3	$x^3 - 3\,x$	6	$He_3/\sqrt{6}$
4	$x^4 - 6\,x^2 + 3$	24	$He_4/\sqrt{24}$
5	$x^5 - 10\,x^3 + 15\,x$	120	$He_5/\sqrt{120}$



Outline

Multivariate polynomials

Tensor product of 1D polynomials

- One defines the multi-indices $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_M\}$, of degree $|\boldsymbol{\alpha}| = \sum_{i=1}^{M} \alpha_i$
- The associated multivariate polynomial reads:

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{i=1}^{M} P^{(i)}_{\alpha_i}(x_i)$$

where $P^{(i)}_{\alpha_i}$ is the univariate polynomial of degree α_i from the orthonormal family associated to variable X_i

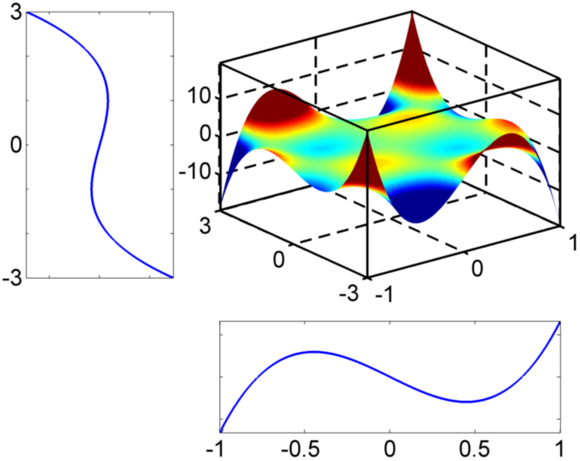
The set of multivariate polynomials $\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}),\ \boldsymbol{\alpha} \in \mathbb{N}^M\}$ forms a basis of $L^2(\Omega, \mathcal{F}, \mathbb{P})$

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

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

Orthonormality of multivariate polynomials

Suppose that the input random vector has independent components:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{M} f_{X_i}(x_i)$$

and consider the tensor product polynomials $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \prod_{i=1}^{M} P_{\alpha_i}^{(i)}(x_i)$ and $\Psi_{\boldsymbol{\beta}}(\boldsymbol{x}) = \prod_{i=1}^{M} P_{\beta_i}^{(i)}(x_i)$. Then:

$$\begin{aligned}
\mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\Psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right] &= \int_{\mathcal{D}_{\boldsymbol{X}}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\Psi_{\boldsymbol{\beta}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_{\mathcal{D}_{\boldsymbol{X}}} \left(\prod_{i=1}^{M} P_{\alpha_i}^{(i)}(x_i) P_{\beta_i}^{(i)}(x_i)\, f_{X_i}(x_i) \right) d\boldsymbol{x} \\
&= \prod_{i=1}^{M} \left(\int_{\mathcal{D}_{X_i}} P_{\alpha_i}^{(i)}(x_i) P_{\beta_i}^{(i)}(x_i)\, f_{X_i}(x_i) dx_i \right) = \prod_{i=1}^{M} \delta_{\alpha_i \beta_i}
\end{aligned}$$

$$\mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\Psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right] = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$$

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: addressing dependence

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

Standard truncation scheme

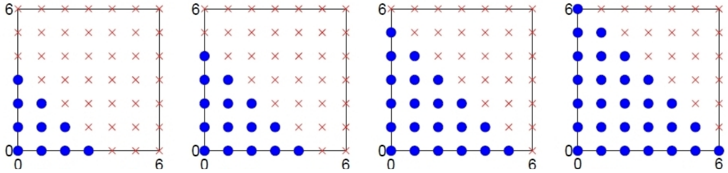
Premise

- The infinite series expansion cannot be handled in pratical computations
- A truncated series must be defined

Standard truncation scheme

Consider all multivariate polynomials of total degree $|\boldsymbol{\alpha}| = \sum_{i=1}^{M} \alpha_i$ less than or equal to p:

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



$|\boldsymbol{\alpha}| \leq 3$ $|\boldsymbol{\alpha}| \leq 4$ $|\boldsymbol{\alpha}| \leq 5$ $|\boldsymbol{\alpha}| \leq 6$

Application example

Computational model $$Y = \mathcal{M}(X_1, X_2)$$

Probabilistic model $$X_1 \sim \mathcal{N}(\mu, \sigma) \quad X_2 \sim \mathcal{U}(a, b)$$

Isoprobabilistic transform
$$X_1 = \mu + \sigma\, \xi_1 \qquad \xi_1 \sim \mathcal{N}(0, 1)$$
$$X_2 = (a + b)/2 + (b - a)\xi_2/2 \qquad \xi_2 \sim \mathcal{U}(-1, 1)$$

Univariate polynomials

- Hermite polynomials in ξ_1, *i.e.* $\tilde{He}_n(\xi_1)$
- Legendre polynomials in ξ_2, *i.e.* $\tilde{P}_n(\xi_2)$

Multivariate polynomials

$$\Psi_{\alpha_1,\alpha_2}(\xi_1, \xi_2) = \tilde{He}_{\alpha_1}(\xi_1) \cdot \tilde{P}_{\alpha_2}(\xi_2)$$

Truncation example

Third order truncation $p = 3$

All the polynomials of ξ_1, ξ_2 that are product of univariate polynomials and whose total degree is less than 3 are considered

j	$\boldsymbol{\alpha}$	$\Psi_{\boldsymbol{\alpha}} \equiv \Psi_j$
0	$[0,0]$	$\Psi_0 = 1$
1	$[1,0]$	$\Psi_1 = \xi_1$
2	$[0,1]$	$\Psi_2 = \sqrt{3}\,\xi_2$
3	$[2,0]$	$\Psi_3 = (\xi_1^2 - 1)/\sqrt{2}$
4	$[1,1]$	$\Psi_4 = \sqrt{3}\,\xi_1\xi_2$
5	$[0,2]$	$\Psi_5 = \sqrt{5/4}\,(3\xi_2^2 - 1)$
6	$[3,0]$	$\Psi_6 = (\xi_1^3 - 3\xi_1)/\sqrt{6}$
7	$[2,1]$	$\Psi_7 = \sqrt{3/2}\,(\xi_1^2 - 1)\xi_2$
8	$[1,2]$	$\Psi_8 = \sqrt{5/4}(3\xi_2^2 - 1)\xi_1$
9	$[0,3]$	$\Psi_9 = \sqrt{7/4}(5\xi_2^3 - 3\xi_2)$

$$\begin{aligned}
\tilde{Y} \equiv \mathcal{M}^{\mathsf{PC}}(\xi_1, \xi_2) &= a_0 + a_1\,\xi_1 + a_2\,\sqrt{3}\,\xi_2 \\
&+ a_3\,(\xi_1^2 - 1)/\sqrt{2} + a_4\,\sqrt{3}\,\xi_1\xi_2 \\
&+ a_5\,\sqrt{5/4}\,(3\xi_2^2 - 1) + a_6\,(\xi_1^3 - 3\xi_1)/\sqrt{6} \\
&+ a_7\,\sqrt{3/2}\,(\xi_1^2 - 1)\xi_2 + a_8\,\sqrt{5/4}(3\xi_2^2 - 1)\xi_1 \\
&+ a_9\,\sqrt{7/4}(5\xi_2^3 - 3\xi_2)
\end{aligned}$$

Conclusions

- Polynomial chaos expansions allow for an intrinsic representation of the random response as a series expansion
- The basis functions are multivariate orthonormal polynomials (based on the input distribution)
- In practice, the input vector is first transformed into independent reduced variables for which classical orthogonal polynomials are well-known
- A truncation scheme shall be introduced for pratical computations, *e.g.* by selecting the maximal degree of the polynomials
- Next step is the computation of the expansion coefficients

Outline

Various methods for computing the coefficients

Intrusive approaches

- Historical approaches: projection of the equations residuals in the Galerkin sense — Ghanem & Spanos, 1991, 2003
- Proper generalized decompositions — Nouy, 2007-2012

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 $\mathcal{X} = \{\boldsymbol{x}^{(i)} \in \mathcal{D}_{\boldsymbol{X}},\ i = 1, \ldots, n\}$ of the input parameters
- Different classes of methods are available:
  - Projection: by simulation or quadrature
  - Stochastic collocation
  - Least-square minimization

Projection

Polynomial chaos expansion

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

By multiplying by $\Psi_{\boldsymbol{\alpha}}$ and taking the expectation one gets:

$$\mathbb{E}\left[Y \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = \sum_{\boldsymbol{\beta} \in \mathbb{N}^M} y_{\boldsymbol{\beta}} \, \overbrace{\mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \, \Psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right]}^{\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}} = y_{\boldsymbol{\alpha}}$$

Estimation techniques

$$y_{\boldsymbol{\alpha}} = \mathbb{E}\left[Y \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}(\boldsymbol{x}) \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \, f_{\boldsymbol{X}}(\boldsymbol{x}) \, d\boldsymbol{x}$$

Computation by full- or Smolyak quadrature

One-dimensional quadrature rules

Consider the following weighted integral, for some positive weight function $w : x \in \mathcal{D}_X \mapsto \mathbb{R}^+$

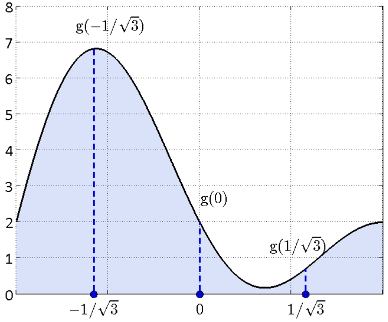
$$\mathcal{I}[g] = \int_{\mathcal{D}_X} g(x)\, w(x)\, dx$$

- A n-point quadrature rule is defined by

$$\mathcal{I}[g] \approx \mathcal{Q}^n[g] \stackrel{\text{def}}{=} \sum_{k=1}^{n} \omega_k\, g(x_k)$$

  where

  - $\{\omega_k,\ k = 1, \ldots, n\}$ are the integration weights
  - $\{x_k,\ k = 1, \ldots, n\}$ are the integration nodes



Gaussian quadrature rules

A Gaussian quadrature rule with n nodes reads:

$$\int_{\mathcal{D}_X} g(x)\, w(x)\, dx \approx \mathcal{Q}^G[g] \stackrel{\text{def}}{=} \sum_{j=1}^{n} \omega_j^G\, g(x_j^G)$$

where:

- The nodes $\left\{ x_j^G,\, j = 1, \, \dots \,, n \right\}$ are the zeros of the n-th orthogonal π_n w.r.t to w

- The weights are given by :

$$\omega_j^G = \frac{< \pi_{n-1} \, , \, \pi_{n-1} >}{\pi'_n(x_j^G).\pi_{n-1}(x_j^G)}$$

The degree of exactness is $d = 2\,n - 1$. It is the largest possible degree of exactness

Multidimensional quadrature

Higher dimensions

Consider the M-dimensional integral: $\mathcal{I}(h) \equiv \int_{D \subset \mathbb{R}^M} h(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$

where $h(.)$ is a function to be integrated against the weight function $f_{\boldsymbol{X}}(.)$:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{X_1}(x_1) \dots f_{X_M}(x_M)$$

The tensorized quadrature scheme consists in replacing each integral by a summation, thus the nested summations:

$$Q^{\boldsymbol{n}}(h) \equiv Q^{(n_1, \dots, n_M)}(h) \equiv \sum_{i_1=1}^{n_1} \cdots \sum_{i_M=1}^{n_M} \omega_{i_1} \cdots \omega_{i_M}\, h(x_{i_1}, \dots, x_{i_M})$$

Computing the integral requires $n_1 \times \cdots \times n_M$ evaluations of the integrand.

Back to the computation of chaos coefficients $(M > 1)$

Each polynomial chaos coefficient $y_{\boldsymbol{\alpha}}$ reads:

$$y_{\boldsymbol{\alpha}} = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}(\boldsymbol{x})\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

- Integrand: $h(\boldsymbol{x}) := \mathcal{M}(\boldsymbol{x})\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$
- Order: $n_i = p + 1, \quad i = 1, \cdots, M$

$$\begin{aligned} \hat{y}_{\alpha} &\equiv Q^{(p+1, \cdots, p+1)}(h) \\ &= \sum_{i_1=1}^{p+1} \cdots \sum_{i_M=1}^{p+1} \omega_{i_1} \cdots \omega_{i_M}\, \mathcal{M}(x_{i_1}, \ldots, x_{i_M})\, \Psi_{\boldsymbol{\alpha}}(x_{i_1}, \ldots, x_{i_M}) \end{aligned}$$

Computational cost : $(p+1)^M$ evaluations of the model

Computational cost

- The cost increases exponentially with M: $N = (p+1)^M$
- Normal industrial (and research!) settings allow at most $\mathcal{O}(100)$ model evaluations
- Industrial problems often use more than 10 variables!
- In some cases, they are very non-linear ($p > 5$)

M	p	N
2	3	16
	5	36
3	3	64
	5	216
5	3	1,024
	5	7,776
10	3	1,048,576
	5	60,466,176

Need for a more efficient scheme in high dimensions

Smolyak quadrature: sparse grids

Smolyak sparse quadrature rule

$$Q^{M,k}_{Smolyak} \equiv \sum_{k+1 \leqslant |\boldsymbol{i}| \leqslant k+M} (-1)^{M+k-|\boldsymbol{i}|} \cdot \binom{M-1}{k+M-|\boldsymbol{i}|} \cdot Q^{\boldsymbol{i}}$$

where:

$$\boldsymbol{i} = i_1, i_2, ..., i_M, \qquad |\boldsymbol{i}| = i_1 + ... + i_M \in \mathbb{N}$$

and

$$Q^{\boldsymbol{i}} = Q^{i_1} \otimes .... \otimes Q^{i_M}$$

Smolyak integration scheme is exact for PC expansions of max. degree p using $k = p$

Outline

Statistical approach: least-square minimization

Berveiller *et al.* (2006)

Principle

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

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{j=0}^{P-1} y_j \, \Psi_j(\mathbf{X}) + \varepsilon_P \equiv \mathbf{Y}^{\mathsf{T}} \boldsymbol{\Psi}(\boldsymbol{X}) + \varepsilon_P$$

where : $\mathbf{Y} = \{y_0, \ldots, y_{P-1}\}$

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

Least-Square Minimization: continuous solution

Least-square minimization

The unknown coefficients are gathered into a vector $\mathbf{Y} = \{y_j, \ j = 0, \dots, P-1\}$, and computed by minimizing the mean square error:

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

Analytical solution (continuous case)

The least-square minimization problem may be solved analytically:

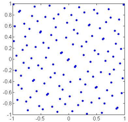
$$\hat{y}_j = \mathbb{E}\left[\mathcal{M}(\boldsymbol{X}) \, \Psi_j(\boldsymbol{X})\right] \qquad \forall \, j = 0, \dots, P-1$$

The solution is identical to the projection solution due to the orthogonality of the regressors

Least-Square Minimization: procedure

$$\hat{\mathbf{Y}} = \arg\min \hat{\mathbb{E}}\left[\left(\mathbf{Y}^{\mathsf{T}}\boldsymbol{\Psi}(\boldsymbol{X}) - \mathcal{M}(\boldsymbol{X})\right)^2\right] = \arg\min_{\boldsymbol{y}\in\mathbb{R}^P} \sum_{i=1}^{n}\left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \sum_{j=0}^{P-1} y_j\,\Psi_j(\boldsymbol{x}^{(i)})\right)^2$$

- Select an experimental design $\mathcal{X} = \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)}\right\}^{\mathsf{T}}$ that covers at best the domain of variation of the parameters



- Evaluate the model response for each sample (exactly as in Monte carlo simulation)

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

- Compute the experimental matrix

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

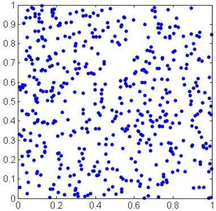
- Solve the least-square minimization problem

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

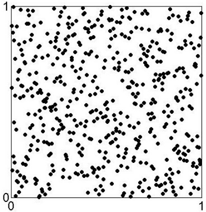
Choice of the experimental design

Random designs

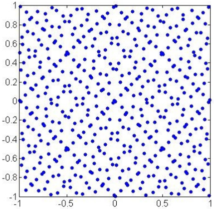
- Monte Carlo samples obtained by standard random generators
- Latin Hypercube designs that are both random and “space-filling”
- Quasi-random sequences (*e.g.* Sobol’ sequence)



Monte Carlo



Latin Hypercube Sampling



Sobol sequence

Size of the experimental design

Size of the ED

The size n of the experimental design shall be scaled with the number of unknown coefficients, *e.g.* $P = \binom{M+p}{p}$

- $n < P$ leads to an underdetermined system
- $n = P$ may lead to overfitting

The thumb rule $n = k\,P$ where $k = 2 - 3$ is used

Error estimators

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

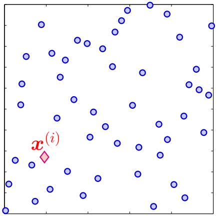
$$E_{gen} = \mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{X}) - \mathcal{M}^{\mathsf{PC}}(\boldsymbol{X})\right)^2\right] \qquad \mathcal{M}^{\mathsf{PC}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

- The empirical error based on the experimental design $\mathcal{X}$ is a poor estimator in case of overfitting

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

Cross-validation techniques

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)

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

- Analytical derivation 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: Ordinary least-squares

```
Input: Computational budget n
Initialization
    Experimental design X = {x^(1), ..., x^(n)}
    Run model Y = {M(x^(1)), ..., M(x^(n))}
PCE construction
    for p = p_min : p_max do
        Select candidate basis A^{M,p}
        Solve OLS problem
        Compute e_LOO(p)
    end
    p* = arg min e_LOO(p)
Return Best PCE of degree p*
```

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

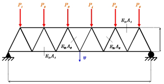
Outline

Curse of dimensionality

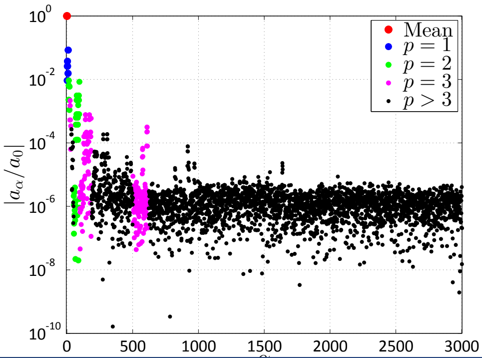
- The cardinality of the truncation scheme $\mathcal{A}^{M,p}$ is $P = \dfrac{(M+p)!}{M!\,p!}$
- Typical computational requirements: $n = OSR \cdot P$ where the oversampling rate is $OSR = 2 - 3$

However ... most coefficients are close to zero !

Example



- Elastic truss structure with $M = 10$ independent input variables
- PCE of degree $p = 5$ ($P = 3,003$ coeff.)



Low-rank truncation schemes

Sparsity-of-effects principle – Ockham's razor

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

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 $\boldsymbol{\alpha}$

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

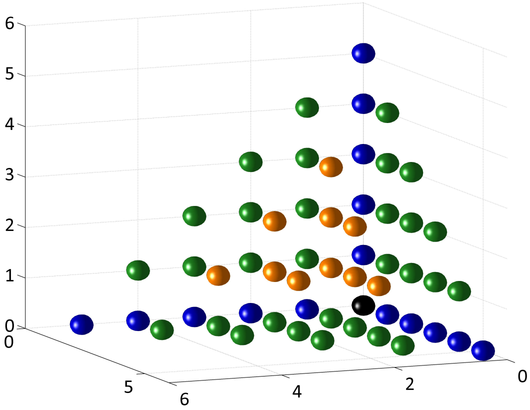
Example: $M = 5, p = 5$, Legendre polynomial chaos

$\boldsymbol{\alpha}$	$\Psi_{\boldsymbol{\alpha}}$	Rank
$[0\ 0\ 0\ 3\ 0]$	$\tilde{P}_3(x_4)$	1
$[2\ 0\ 0\ 0\ 1]$	$\tilde{P}_2(x_1) \cdot \tilde{P}_1(x_5)$	2
$[1\ 1\ 2\ 0\ 1]$	$\tilde{P}_1(x_1) \cdot \tilde{P}_1(x_2) \cdot \tilde{P}_2(x_3) \cdot \tilde{P}_1(x_5)$	4

Low-rank truncation set

Definition

$$\mathcal{A}^{M,p,r} = \{\boldsymbol{\alpha} \in \mathbb{N}^M \ : \ |\boldsymbol{\alpha}| \leq p, ||\boldsymbol{\alpha}||_0 \leq r\} \qquad r \leq p \ , \ r \leq M$$



All ranks ≤ 3

Hyperbolic truncation sets

Sparsity-of-effects principle

Blatman & Sudret, Prob. Eng. Mech (2010); J. Comp. Phys (2011)

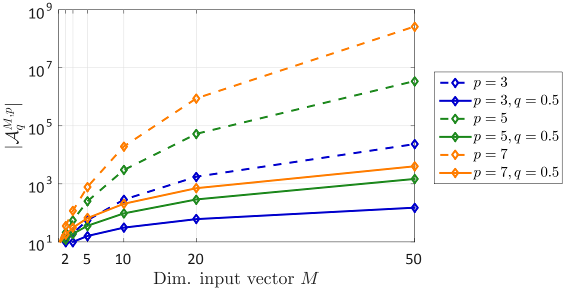
In most engineering problems, only low-order interactions between the input variables are relevant

- $q-$norm of a multi-index $\boldsymbol{\alpha}$:

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

- Hyperbolic truncation sets:

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



Compressive sensing approaches

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

- 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, orthogonal matching pursuit, Bayesian compressive sensing

Least Angle Regression

Efron *et al.* (2004)
Blatman & Sudret (2011)

- Least Angle Regression (LAR) solves the LASSO problem for different values of the penalty constant in a single run without matrix inversion
- Leave-one-out cross validation error allows one to select the best model

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

Outline

Statistical moments

From the orthonormality of the polynomial chaos basis one gets:

$$\mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = 0 \qquad \text{for } \boldsymbol{\alpha} \neq \mathbf{0}$$
$$\mathbb{E}\left[\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\Psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right] = 0 \qquad \text{for } \boldsymbol{\alpha} \neq \boldsymbol{\beta}$$

Mean value

$\hat{\mu}_Y = y_0$ — The mean value is the first coefficient of the series

Variance

$$\hat{\sigma}_Y^2 \stackrel{\text{def}}{=} \mathbb{E}\left[\left(Y^{PC} - \hat{\mu}_Y\right)^2\right] = \mathbb{E}\left[\left(\sum_{\boldsymbol{\alpha}\in\mathcal{A}\setminus\mathbf{0}} y_{\boldsymbol{\alpha}}\, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right]$$

$\hat{\sigma}_Y^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{A}\setminus\mathbf{0}} y_{\boldsymbol{\alpha}}^2$ — The variance is computed as the sum of the squares of the remaining coefficients

Probability density function

Principle

The polynomial series expansion may be used as a stochastic response surface

- A large sample set $\boldsymbol{\xi}$ of reduced variables is drawn, say of size $n_{sim} = 10^5 - 10^6$:

$$\mathcal{X}_{sim} = \{\boldsymbol{\xi}_j,\ j = 1, \ldots, n_{sim}\}$$

- The truncated series is evaluated onto this sample:

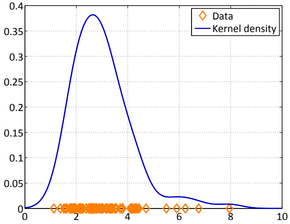
$$\mathcal{Y}_{sim} = \left\{ \mathfrak{y}_j = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_j),\ j = 1, \ldots, n_{sim} \right\}$$

- The obtained sample set is plotted using histograms or kernel density smoothing

Probability density function

Kernel smoothing

$$\hat{f}_Y(y) = \frac{1}{n_{sim}\, h} \sum_{j=1}^{n_{sim}} K\left(\frac{y - \mathfrak{y}_j}{h}\right)$$



- Kernel function : $K(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$
- Bandwidth:

$$h = 0.9\, {n_{sim}}^{-1/5}\, \min\left(\hat{\sigma}_{\mathcal{Y}}, (Q_{0.75} - Q_{0.25})/1.34\right)$$

where $Q_{0.75} - Q_{0.25}$ is the empirical inter-quartile computed from the sample

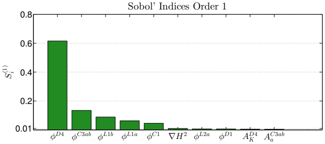
Step C': sensitivity analysis

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)

- Screening: detect input parameters whose uncertainty has no impact on the output variability
- Feature setting: detect input parameters which allow one to best decrease the output variability when set to a deterministic value
- Exploration: detect interactions between parameters, *i.e.* joint effects not detected when varying parameters one-at-a-time



Variance decomposition (Sobol' indices)

Sensitivity analysis

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

  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, CSM (2006); RESS (2008)

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

Interaction sets

For a given $\mathbf{u} \stackrel{\text{def}}{=} \{i_1, \dots, i_s\}$: $\quad \mathcal{A}_{\mathbf{u}} = \{\boldsymbol{\alpha} \in \mathcal{A} \, : \, k \in \mathbf{u} \Leftrightarrow \alpha_k \neq 0\}$

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

PC-based Sobol' indices

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

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

Example

Computational model $Y = \mathcal{M}(X_1, X_2)$

Probabilistic model $X_i \sim \mathcal{N}(\mu_i, \sigma_i)$

Isoprobabilistic transform $X_i = \mu_i + \sigma_i \, \xi_i$

Chaos degree $p = 3$, *i.e.* $P = 10$ terms

j	$\boldsymbol{\alpha}$	$\Psi_{\boldsymbol{\alpha}} \equiv \Psi_j$
0	$[0, 0]$	$\Psi_0 = 1$
1	$[1, 0]$	$\Psi_1 = \xi_1$
2	$[0, 1]$	$\Psi_2 = \xi_2$
3	$[2, 0]$	$\Psi_3 = (\xi_1^2 - 1)/\sqrt{2}$
4	$[1, 1]$	$\Psi_4 = \xi_1 \xi_2$
5	$[0, 2]$	$\Psi_5 = (\xi_2^2 - 1)/\sqrt{2}$
6	$[3, 0]$	$\Psi_6 = (\xi_1^3 - 3\xi_1)/\sqrt{6}$
7	$[2, 1]$	$\Psi_7 = (\xi_1^2 - 1)\xi_2/\sqrt{2}$
8	$[1, 2]$	$\Psi_8 = (\xi_2^2 - 1)\xi_1/\sqrt{2}$
9	$[0, 3]$	$\Psi_9 = (\xi_2^3 - 3\xi_2)/\sqrt{6}$

Variance

$$D = \sum_{j=1}^{9} a_j^2$$

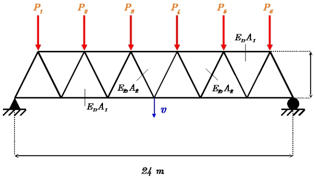
Sobol' indices

$$\begin{aligned} S_1 &= \left(a_1^2 + a_3^2 + a_6^2\right)/D \\ S_2 &= \left(a_2^2 + a_5^2 + a_9^2\right)/D \\ S_{12} &= \left(a_4^2 + a_7^2 + a_8^2\right)/D \end{aligned}$$

Outline

Elastic truss structures

Blatman *et al.* , 2007



10 independent input variables

- 4 describing the bars properties
- 6 describing the loads

Questions

PDF of the max. deflection, statistical moments, probability of failure

$$V = \mathcal{M}^{\mathsf{FE}} (E_1, E_2, A_1, A_2, P_1, \ldots, P_6)$$

Probabilistic model

Parameters	Name	Distribution	Mean	Std. Deviation
Young's modulus	E_1, E_2 (Pa)	Lognormal	2.10×10^{11}	2.10×10^{10}
Hor. bars section	A_1 (m^2)	Lognormal	2.0×10^{-3}	2.0×10^{-4}
Vert. bars section	A_2 (m^2)	Lognormal	1.0×10^{-3}	1.0×10^{-4}
Loads	P_1-P_6 (N)	Gumbel	5.0×10^4	7.5×10^3

Isoprobabilistic transform

Lognormal and Gumbel distributions are transformed into reduced Gaussian variables

- Lognormal variables E_1, E_2, A_1, A_2

$$X_i \sim \mathcal{LN}(\lambda_i, \zeta_i)$$
$$X_i = e^{\lambda_i + \zeta_i \, U_i} \qquad U_i \sim \mathcal{N}(0,1)$$

- Gumbel variables $P_1, \ldots, P_6$

$$P_j \sim \mathcal{G}(\mu_j, \beta_j) \qquad F_{P_j}(x) = \exp\left[-\exp\left[-(x-\mu_j)/\beta_j\right]\right]$$

Thus:

$$P_j = \mu_j - \beta_j \ln\left(-\ln \Phi(U_j)\right) \qquad U_j \sim \mathcal{N}(0,1)$$

where the parameters (μ_j, β_j) are linked to the moments by:

$$\mathbb{E}\left[P_j\right] = \mu_j + 0.577216 \, \beta_j \qquad \sigma_{P_j} = \frac{\pi \beta_j}{\sqrt{6}}$$

Full polynomial chaos expansions

Comparison of methods

Case	Degree p	$P = \|\mathcal{A}\|$	Method	Cost
A	2	66	Quadrature	59,049
B	2	66	Sparse quadrature	231
C_1	2	66	Least-squares ($n = 2\,P$)	132
C_2	2	66	Least-squares ($n = 3\,P$)	198
D	3	286	Quadrature	1,048,576
E	3	286	Smolyak quadrature	1,771
F_1	3	286	Least-squares ($n = 2\,P$)	572
F_2	3	286	Least-squares ($n = 3\,P$)	858
G	4	1,001	Smolyak quadrature	10626
H_1	4	1,001	Least-squares ($n = 2\,P$)	2,002
H_2	4	1,001	Least-squares ($n = 3\,P$)	3,003

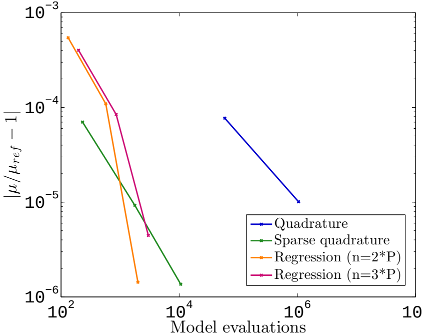
Full polynomial chaos expansions

Moments and quantiles of the maximal deflection (in cm)

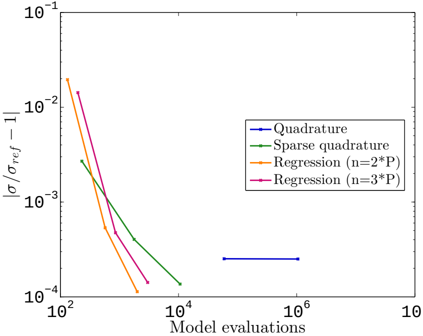
Case	μ_V	σ_V	$v_{95\%}$	$v_{99\%}$	Cost
Reference	7.9400	1.1078	9.9012	10.9241	1,000,000
A	7.9407	1.1079	9.9015	10.8854	59,049
B	7.9406	1.1052	9.8883	10.8554	231
C_1	7.9357	1.0865	9.8485	10.7939	132
C_2	7.9369	1.0924	9.8604	10.8118	198
D	7.9400	1.1085	9.9030	10.9237	1,048,576
E	7.9400	1.1086	9.9037	10.9232	1,771
F_1	7.9392	1.1076	9.8987	10.9149	572
F_2	7.9394	1.1077	9.8991	10.9152	858
G	7.9401	1.1083	9.9006	10.9248	10,626
H_1	7.9401	1.1083	9.9013	10.9236	2,002
H_2	7.9401	1.1083	9.9014	10.9239	3,003

NB: Reference values are obtained from $n = 10^6$ points (Sobol' sequence)

Full PCE: Convergence curves



Mean value



Standard deviation

Sparse polynomial chaos expansions

Set up

- The size of the experimental design is fixed to $n = 50,\ 100,\ 200,\ 500,\ 1000,\ 2000$. Sobol points are used
- The standard truncation scheme is used $(q = 1)$. Different candidate sets $\mathcal{A}^{10,p}$ are used with $p = 2,\ 3,\ \dots,10$.
- The best sparse expansion is retained by cross-validation

Results

n	p_{opt}	$\lvert\mathcal{A}^{10,p_{opt}}\rvert$	# Terms	Index of sparsity	ϵ_{LOO}
50	3	286	6	0.0210	2.9384e-01
100	2	66	49	0.7424	3.4961e-03
200	3	286	81	0.2832	1.6448e-03
500	3	286	151	0.5280	4.9202e-05
1000	4	1001	381	0.3806	7.0862e-06
2000	5	3003	473	0.1575	1.9126e-06

Sparse polynomial chaos expansions

Set up

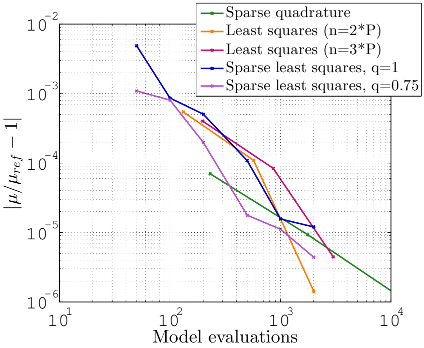
- The same calculations are carried out using *a priori* a hyperbolic truncation set with $q-$norm $q = 0.75$

Results

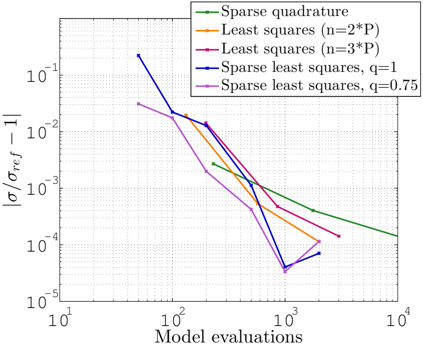
n	p_{opt}	$\lvert\mathcal{A}^{10,p_{opt}}\rvert$	# Terms	Index of sparsity	ϵ_{LOO}
50	2	66	19	0.2879	4.2476e-02
100	3	286	54	0.1888	3.8984e-03
200	4	1001	121	0.1209	3.9940e-04
500	5	3003	279	0.0929	4.0975e-05
1000	6	8008	579	0.0723	2.2236e-06
2000	7	19448	806	0.0414	1.5752e-07

- Higher degree terms are included (more sparsity)
- Better accuracy as measured by ϵ_{LOO}

Sparse PCE: Convergence curves



Mean value



Standard deviation

Structural reliability analysis

Limit state function: $g(\boldsymbol{X}) \equiv v_{\max} - \mathcal{M}(E_1, E_2, A_1, A_2, P_1, \dots, P_6)$

Full PCE

$p = 3$, Smolyak quadrature (1,771 runs)

Threshold $v_{\max}$ (cm)	Reference		Smolyak quadrature	
	P_f	β	P_f	β
10	$4.31\ 10^{-2}$	1.71	$4.29\ 10^{-2}$	1.71
11	$8.70\ 10^{-3}$	2.37	$8.70\ 10^{-3}$	2.37
12	$1.50\ 10^{-3}$	2.96	$1.50\ 10^{-3}$	2.97
14	$3.49\ 10^{-5}$	3.97	$2.83\ 10^{-5}$	4.02
16	$6.03\ 10^{-7}$	4.85	$4.01\ 10^{-7}$	4.93

† $\beta = -\Phi^{-1}(P_f)$

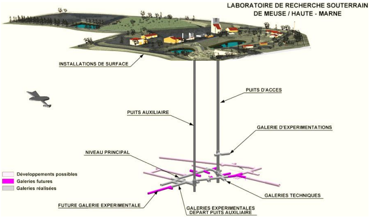
Sparse PC

LAR (500 runs)

	Reference (10^5 runs)	LAR (500 runs)
10 cm	4.39e-02 ± 3.0%	4.30e-02 ± 0.9%
11 cm	8.61e-03 ± 6.7%	8.71e-03 ± 2.1%
12 cm	1.62e-03 ± 15.4%	1.51e-03 ± 5.1%
13 cm	2.20e-04 ± 41.8%	2.03e-04 ± 13.8%

Outline

Example: sensitivity analysis in hydrogeology



Source: http://www.futura-sciences.com/

Source: http://lexpansion.lexpress.fr/

- When assessing a nuclear waste repository, 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 system

- Computational models have numerous input parameters (in each geological layer) that are difficult to measure, and that show scattering

Geological model

Joint work with University of Neuchâtel

Deman, Konakli, Sudret, Kerrou, Perrochet & Benabderrahmane, Reliab. Eng. Sys. Safety (2016)

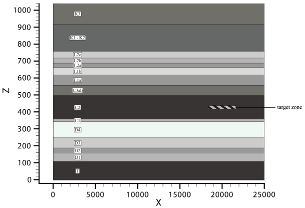
- Two-dimensional idealized model of the Paris Basin (25 km long / 1,040 m depth) with 5×5 m mesh (10^6 elements)
- Steady-state flow simulation with Dirichlet boundary conditions:

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

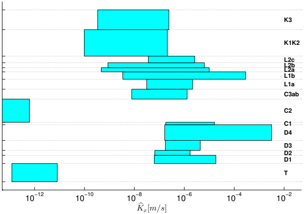
- 15 homogeneous layers with uncertainties in:
  - Porosity (resp. hydraulic conductivity)
  - Anisotropy of the layer properties (inc. dispersivity)
  - Boundary conditions (hydraulic gradients)

78 input parameters

Sensitivity analysis



Geometry of the layers



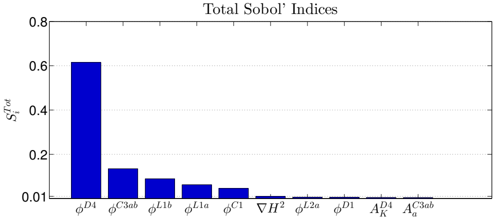
Conductivity of the layers

Question

What are the parameters (out of 78) whose uncertainty drives the uncertainty of the prediction of the mean life-time expectancy?

Sensitivity analysis: results

Technique: Sobol'indices computed from polynomial chaos expansions



Parameter	$\sum_j S_j$
ϕ (resp. K_x)	0.8664
A_K	0.0088
θ	0.0029
α_L	0.0076
A_α	0.0000
∇H	0.0057

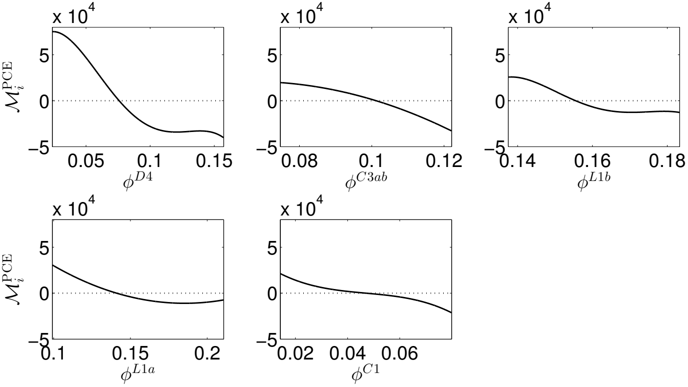
Conclusions

- Only 200 model runs allow one to detect the 10 important parameters out of 78
- Uncertainty in the porosity/conductivity of 5 layers explain 86% of the variability
- Small interactions between parameters detected

Bonus: univariate effects

The univariate effects of each variable are obtained as a straightforward post-processing of the PCE

$$\mathcal{M}_i(x_i) \stackrel{\text{def}}{=} \mathbb{E}\left[\mathcal{M}(\boldsymbol{X}) | X_i = x_i\right], \; i = 1, \, \dots \, , M$$

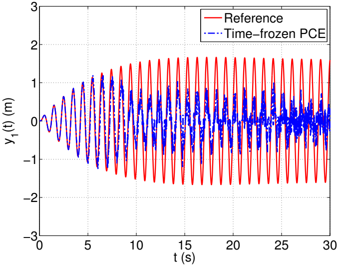


Polynomial chaos expansions in structural dynamics

Spiridonakos et al. (2015); Mai & Sudret, ICASP'2015; Mai *et al.* , 2016

Premise

- For dynamical systems, the complexity of the map $\boldsymbol{x} \mapsto \mathcal{M}(\boldsymbol{x}, t)$ increases with time.
- Time-frozen PCE does not work beyond first time instants



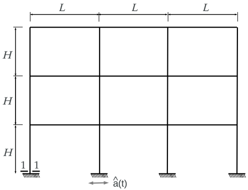
- Use of non linear autoregressive with exogenous input models (NARX) to capture the dynamics:

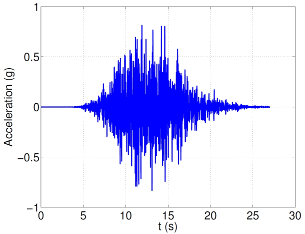
$$\begin{aligned} y(t) = &\mathcal{F}\left(x(t), \ldots, x(t-n_x),\right. \\ &\left. y(t-1), \ldots, y(t-n_y)\right) + \epsilon_t \end{aligned}$$

- Expand the NARX coefficients of different random trajectories onto a PCE basis

$$y(t, \boldsymbol{\xi}) = \sum_{i=1}^{n_g} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i} \vartheta_{i,\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \, g_i(\boldsymbol{z}(t)) + \epsilon(t, \boldsymbol{\xi})$$

Application: earthquake engineering





Rezaeian & Der Kiureghian (2010)

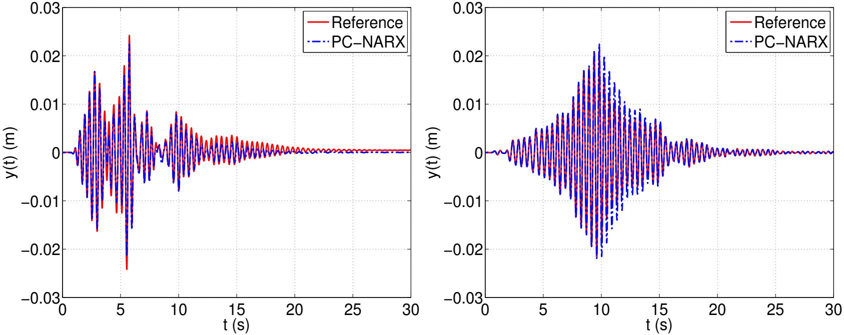
- 2D steel frame with uncertain properties submitted to synthetic ground motions
- Experimental design of size 300

- Ground motions obtained from modulated, filtered white noise

$$x(t) = q(t, \boldsymbol{\alpha}) \sum_{i=1}^{n} s_i \left(t, \boldsymbol{\lambda}(t_i)\right) \cdot \xi_i \qquad \xi_i \sim \mathcal{N}(0,1)$$

Application: earthquake engineering

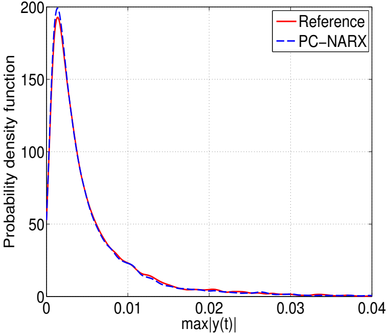
Surrogate model of single trajectories



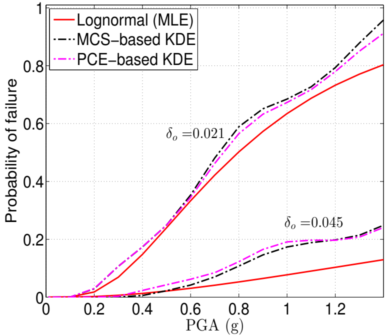
Application: earthquake engineering

First-storey drift

- PC-NARX calibrated based on 300 simulations
- Reference results obtained from 10,000 Monte Carlo simulations



PDF of max. drift

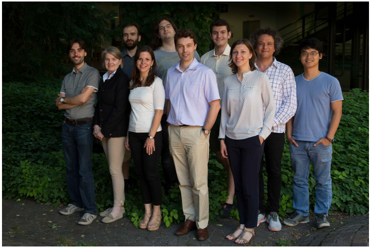


Fragility curves for two drift thresholds

Conclusions

- Uncertainty quantification for engineering applications require non-intrusive, parsimonious techniques to solve propagation, sensitivity and reliability problems
- Polynomial chaos expansions represent the quantities of interest as a multivariate orthonormal polynomial series in the input variables
- Coefficients can be computed by projection, ordinary least-square and compressive sensing
- Post-processing of the series provides moments, quantiles, PDF, probabilities of failure, sensitivity indices, etc.
- Problems involving $\mathcal{O}(10)$ input variables can be solved in $\mathcal{O}(100)$ runs

Questions ?



Chair of Risk, Safety & Uncertainty Quantification

www.rsuq.ethz.ch

The Uncertainty Quantification Laboratory

www.uqlab.com

Thank you very much for your attention !