Sparse polynomial chaos expansions and application to sensitivity analysis
BOQUSE’2013 – December 18th – INRIA

Author(s):
Sudret, Bruno

Publication Date:
2013

Permanent Link:
https://doi.org/10.3929/ethz-a-010060789

Rights / License:
In Copyright - Non-Commercial Use Permitted
Sparse polynomial chaos expansions and application to sensitivity analysis

B. Sudret

Chair of Risk, Safety & Uncertainty Quantification

BOQUSE’2013 – December 18th – INRIA
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.
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
- Advanced meta-models (kriging, support vector machines)
- Bayesian model calibration and stochastic inverse problems
- Global sensitivity analysis
- Reliability-based design optimization

http://www.rsuq.ethz.ch
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
- Advanced meta-models (kriging, support vector machines)
- Bayesian model calibration and stochastic inverse problems
- **Global sensitivity analysis**
- Reliability-based design optimization

http://www.rsuq.ethz.ch
Uncertainty quantification in industrial applications

- Uncertainty quantification arrives on top of well defined simulation procedures (legacy codes)

- The computational models are complex: coupled problems (thermo-mechanics), plasticity, large strains, contact, buckling, etc.

- A single simulation is already costly (e.g. several hours)

- Engineers focus on so-called quantities of interest, e.g. maximum displacement, average stress, etc.
Uncertainty quantification in industrial applications

- The input variables modelling aleatory uncertainty are often non Gaussian.
- The size of the input random vector is typically 10-100.
- UQ procedures shall be sufficiently general to be applied with little adaptation to a variety of problems.

Need for non intrusive and parsimonious methods for uncertainty quantification.
Global framework for uncertainty quantification

Step A
Model(s) of the system
Assessment criteria

Step B
Quantification of sources of uncertainty

Step C
Uncertainty propagation

Computational model

Random variables

Step C'
Sensitivity analysis

Global framework for uncertainty quantification

Step A
Model(s) of the system
Assessment criteria

Step B
Quantification of sources of uncertainty

Step C
Uncertainty propagation

Random variables
Computational model

Step C'
Sensitivity analysis

Global framework for uncertainty quantification

**Step A**
Model(s) of the system
Assessment criteria

**Step B**
Quantification of sources of uncertainty

**Step C**
Uncertainty propagation

**Step C’**
Sensitivity analysis

Random variables
Computational model
Distribution
Mean, std. deviation
Probability of failure

Global framework for uncertainty quantification

Step A
Model(s) of the system
Assessment criteria

Step B
Quantification of sources of uncertainty

Step C
Uncertainty propagation

Step C’
Sensitivity analysis

Random variables

Computational model

Distribution
Mean, std. deviation
Probability of failure

Outline

1. Sparse polynomial chaos scheme
   - Sparse truncation schemes
   - Computation of the coefficients
   - Error estimation and validation

2. Adaptive algorithms for sparse expansions

3. Sensitivity analysis
   - Sobol’ indices
   - Case of dependent inputs: ANCOVA

4. Application examples in sensitivity analysis
   - Ishigami function
   - Morris function
   - Bending beam
Spectral approach

- The input parameters are modelled by a random vector $\mathbf{X}$ over a probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(d\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x}$.

- The response random vector $\mathbf{Y} = \mathcal{M}(\mathbf{X})$ is considered as an element of $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

- A basis of multivariate orthogonal polynomials is built up with respect to the input PDF (assuming independent components).

- The response random vector $\mathbf{Y}$ is completely determined by its coordinates in this basis.
Spectral approach

- The input parameters are modelled by a random vector $X$ over a probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(dx) = f_X(x) \, dx$

- The response random vector $Y = \mathcal{M}(X)$ is considered as an element of $L^2(\Omega, \mathcal{F}, \mathbb{P})$

- A basis of multivariate orthogonal polynomials is built up with respect to the input PDF (assuming independent components)

- The response random vector $Y$ is completely determined by its coordinates in this basis
Spectral approach

- The input parameters are modelled by a random vector $X$ over a probabilistic space $(\Omega, \mathcal{F}, P)$ such that $P(dx) = f_X(x) \, dx$

- The response random vector $Y = M(X)$ is considered as an element of $L^2(\Omega, \mathcal{F}, P)$

- A basis of multivariate orthogonal polynomials is built up with respect to the input PDF (assuming independent components)

- The response random vector $Y$ is completely determined by its coordinates in this basis

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

where:

- $y_{\alpha}$: coefficients to be computed (coordinates)
- $\Psi_{\alpha}(X)$: basis
Univariate orthogonal polynomials

For each marginal distribution $f_{X_i}(x_i)$ one can define a functional inner product:

$$\langle \phi_1, \phi_2 \rangle_i = \int_{D_i} \phi_1(x) \phi_2(x) f_{X_i}(x_i) \, dx_i$$

and a family of orthogonal polynomials $\{P^{(i)}_k, \, k \in \mathbb{N}\}$ such that:

$$\langle P^{(i)}_j, P^{(i)}_k \rangle = \int P^{(i)}_j(x) P^{(i)}_k(x) f_{X_i}(x) \, dx = a^i_j \delta_{jk}$$

Classical families

<table>
<thead>
<tr>
<th>Type of variable</th>
<th>Weight function</th>
<th>Orthogonal polynomials</th>
<th>Hilbertian basis $\psi_k(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>$1_{]-1,1[}(x)/2$</td>
<td>Legendre $P_k(x)$</td>
<td>$P_k(x)/\sqrt{2k+1}$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>$1/\sqrt{2\pi} , e^{-x^2}/2$</td>
<td>Hermite $H_{\ell_k}(x)$</td>
<td>$H_{\ell_k}(x)/\sqrt{k!}$</td>
</tr>
<tr>
<td>Gamma</td>
<td>$x^a e^{-x} , 1_{\mathbb{R}^+}(x)$</td>
<td>Laguerre $L^a_k(x)$</td>
<td>$L^a_k(x)/\sqrt{\Gamma(k+a+1)/k!}$</td>
</tr>
<tr>
<td>Beta</td>
<td>$1_{]-1,1[}(x) , (1-x)^a(1+x)^b , B(a+b)/(B(a)B(b))$</td>
<td>Jacobi $J_{a,b}^k(x)$</td>
<td>$J_{a,b}^k(x)/\sqrt{2^{a+b+1}\Gamma(k+a+1)\Gamma(k+b+1)\Gamma(k+a+b+1)\Gamma(k+1)}$</td>
</tr>
</tbody>
</table>

Xiu & Karniadakis (2002)
Polynomial chaos basis

Univariate orthogonal polynomials

For each marginal distribution $f_{X_i}(x_i)$ one can define a functional inner product:

$$\langle \phi_1, \phi_2 \rangle_i = \int_{\mathcal{D}_i} \phi_1(x) \phi_2(x) f_{X_i}(x_i) \, dx_i$$

and a family of orthogonal polynomials $\{P^{(i)}_k, \, k \in \mathbb{N}\}$ such that:

$$\langle P^{(i)}_j, P^{(i)}_k \rangle = \int P^{(i)}_j(x) P^{(i)}_k(x) f_{X_i}(x) \, dx = a_{ij} \delta_{jk}$$

Classical families

<table>
<thead>
<tr>
<th>Type of variable</th>
<th>Weight function</th>
<th>Orthogonal polynomials</th>
<th>Hilbertian basis $\psi_k(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>$1 - 1, 1[1(x)/2$</td>
<td>Legendre $P_k(x)$</td>
<td>$P_k(x)/\sqrt{\frac{1}{2k+1}}$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>$\frac{1}{\sqrt{2\pi}} e^{-x^2}/2$</td>
<td>Hermite $H_{\varepsilon_k}(x)$</td>
<td>$H_{\varepsilon_k}(x)/\sqrt{k!}$</td>
</tr>
<tr>
<td>Gamma</td>
<td>$x^a e^{-x} , 1_{\mathbb{R}^+}(x)$</td>
<td>Laguerre $L^a_k(x)$</td>
<td>$L^a_k(x)/\sqrt{\frac{\Gamma(k+a+1)}{k!}}$</td>
</tr>
<tr>
<td>Beta</td>
<td>$1_{-1, 1[1(x)}/2$</td>
<td>Jacobi $J^{a,b}_k(x)$</td>
<td>$J^{a,b}_k(x)/\sqrt{\frac{\Gamma(k+a+1)\Gamma(k+b+1)}{\Gamma(k+a+b+1)\Gamma(k+1)}}$</td>
</tr>
</tbody>
</table>

Xiu & Karniadakis (2002)
Multivariate polynomials

Let us define the multi-indices (tuples) $\alpha = \{\alpha_1, \ldots, \alpha_M\}$, of degree $|\alpha| = \sum_{i=1}^{M} \alpha_i$. The associated multivariate polynomial reads:

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

The set of multivariate polynomials $\{\Psi_\alpha, \alpha \in \mathbb{N}^M\}$ forms a basis of the space of second order variables:

$$Y = \sum_{\alpha \in \mathbb{N}^M} y_\alpha \Psi_\alpha(X)$$
Curse of dimensionality

Truncated series

- A truncation scheme is selected and the associated finite set of multi-indices is generated
- The common truncation scheme considers all polynomials up to a given total degree, e.g.:

\[ \mathcal{A}^{M,p} = \{ \alpha \in \mathbb{N}^M : |\alpha| \leq p \} \]

**Drawback:** The number of unknown coefficients \( \text{card} \mathcal{A}^{M,p} \) grows polynomially both in \( M \) and \( p \):

\[ \text{card} \mathcal{A}^{M,p} = \binom{M + p}{p} \]

"Full" expansions are not tractable when \( M \geq 10 \)

**Solutions:**
- Sparse truncation schemes, based on the sparsity-of-effect principle
- **Adaptive** algorithms for the construction of the PC expansion
Why are sparse representations relevant?

Blatman & S., (2008); Cohen, Schwab et al. (2010-13); Doostan & Owhadi (2011)

**Sparsity-of-effects principle:** in usual problems, only low-order interactions between the input variables are relevant. One shall select PC approximations using *low-rank* monomials

**Degree** of a multi-index \( \alpha \): total degree of polynomial \( \Psi_\alpha \)

\[
|\alpha| \equiv \|\alpha\|_1 = \sum_{i=1}^{M} \alpha_i
\]

**Rank** of a multi-index \( \alpha \): number of active variables of \( \Psi_\alpha \) (non zero terms of multi-index \( \alpha \))

\[
\|\alpha\|_0 = \sum_{i=1}^{M} 1_{\{\alpha_i > 0\}}
\]
Two selection techniques

- Low-rank index sets:
  \[ \mathcal{A}^{M,p,j} = \{ \alpha \in \mathbb{N}^M : |\alpha| \leq p, ||\alpha||_0 \leq j \} \]

- Hyperbolic sets:
  \[ \mathcal{A}_{q,p}^{M} = \{ \alpha \in \mathbb{N}^M : ||\alpha||_q \leq p \} \]
  where \[ ||\alpha||_q \equiv \left( \sum_{i=1}^{M} \alpha_i^q \right)^{1/q} , \quad 0 < q < 1 \]

Limit cases

- \( q = 1 \) : common truncation scheme (all polynomials of maximal total degree \( p \))
- \( q \rightarrow 0 \) : additive model (no interaction)
The hyperbolic norm primarily selects the high-degree polynomials in one single variable and then the polynomials involving few interaction.
**Index of sparsity**

- **Common truncation** \( \mathcal{A}^{M,p} \)
- **Hyperbolic truncation** \( \mathcal{A}_q^{M,p} \)

**Effect of the hyperbolic truncation scheme**

\[
IS_1 = \frac{\text{card } \mathcal{A}_q^{M,p}}{\text{card } \mathcal{A}^{M,p}}
\]
Index of sparsity

Common truncation ($\mathcal{A}^M,p$)

Hyperbolic truncation ($\mathcal{A}^M_q,p$)

- Effect of the hyperbolic truncation scheme

$$IS1 = \frac{\text{card } \mathcal{A}^M_q}{\text{card } \mathcal{A}^M,p}$$
Outline

1. Sparse polynomial chaos scheme
   - Sparse truncation schemes
   - Computation of the coefficients
   - Error estimation and validation

2. Adaptive algorithms for sparse expansions

3. Sensitivity analysis

4. Application examples in sensitivity analysis
Various methods for computing the coefficients

Intrusive approaches

- Historical approaches: projection of the equations residuals in the Galerkin sense
  
  \textit{Ghanem et al.; Le Maître et al., Babuska, Tempone et al.; Karniadakis et al., etc.}

- Proper generalized decompositions
  
  \textit{Nouy et al., 2007-10}

Non intrusive approaches

- Non intrusive methods consider the computational model $\mathcal{M}$ as a black box

  \textit{They rely upon a design of numerical experiments, i.e. a $n$-sample $\mathcal{X} = \{x^{(i)} \in \mathcal{D}_X, i = 1, \ldots, n\}$ of the input parameters}

- Different classes of methods are available:
  
  - \textbf{projection: by simulation or quadrature}
    \textit{Matthies & Keese, 2005; Le Maître et al.}
  
  - \textbf{stochastic collocation}
    \textit{Xiu, 2007-09; Nobile, Tempone et al., 2008; Ma & Zabaras, 2009}
  
  - \textbf{least-square minimization}
    \textit{Berveiller et al., 2006; Blatman & S., 2008-11}
**Principle**

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

\[
Y = \mathcal{M}(X) = \sum_{j=0}^{P-1} y_j \Psi_j(X) + \varepsilon_P \equiv Y^T \Psi(X) + \varepsilon_P
\]

where: \(Y = \{y_0, \ldots, y_{P-1}\}\)

\[
\Psi(x) = \{\Psi_0(x), \ldots, \Psi_{P-1}(x)\}
\]
Least-Square Minimization: continuous solution

Least-square minimization

The coefficients are gathered into a vector $\hat{Y}$, and computed by minimizing the mean square error:

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

Analytical solution (continuous case)

The least-square minimization problem may be solved analytically:

$$\hat{Y} = E \left[ M(X) \Psi(X) \right]$$

The solution is identical to the projection solution due to the orthogonality properties of the regressors.
Resolution

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

\[
\hat{Y}^{L.S} = \arg \min \hat{E} \left[ \left( Y^T \Psi(X) - M(X) \right)^2 \right]
\]

\[
= \arg \min \frac{1}{n} \sum_{i=1}^{n} \left( Y^T \Psi(x^{(i)}) - M(x^{(i)}) \right)^2
\]
Least-Square Minimization: discretized solution

- Select an experimental design
  \[ \mathcal{X} = \{ \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)} \}^T \]
  that covers at best the domain of variation of the parameters

- Evaluate the model response for each sample (exactly as in Monte Carlo simulation)
  \[ \mathcal{M} = \{ \mathcal{M}(\mathbf{x}^{(1)}), \ldots, \mathcal{M}(\mathbf{x}^{(n)}) \}^T \]

- Compute the experimental matrix
  \[ A_{ij} = \Psi_j (\mathbf{x}^{(i)}) \quad i = 1, \ldots, n \quad j = 0, \ldots, P - 1 \]

- Solve the least-square minimization problem
  \[ \hat{\mathbf{Y}} = (A^T A)^{-1} A^T \mathcal{M} \]
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)
Outline

1. Sparse polynomial chaos scheme
   - Sparse truncation schemes
   - Computation of the coefficients
   - Error estimation and validation

2. Adaptive algorithms for sparse expansions

3. Sensitivity analysis

4. Application examples in sensitivity analysis
Validation of the surrogate model

- The truncated series expansions are convergent in the mean square sense. However one does not know in advance where to truncate (problem-dependent).
- Usually the series is truncated according to the total maximal degree of the polynomials, say \( p = 2, 3, 4 \) etc.
- The recent research deals with the development of error estimates:
  - adaptive integration in the projection approach
  - cross validation in the least-squares approach
The least-squares technique is based on the minimization of the mean square error. The generalization error is defined as:

$$E_{\text{gen}} = \mathbb{E} \left[ (\mathcal{M}(\mathbf{X}) - \mathcal{M}^{\text{PC}}(\mathbf{X}))^2 \right]$$

It may be estimated by the empirical error using the already computed response quantities:

$$E_{\text{emp}} = \frac{1}{n} \sum_{i=1}^{n} (\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PC}}(\mathbf{x}^{(i)}))^2$$

The coefficient of determination $R^2$ is often used as an error estimator:

$$R^2 = 1 - \frac{E_{\text{emp}}}{\hat{\mathbb{V}}[\mathcal{Y}]} \quad \hat{\mathbb{V}}[\mathcal{Y}] = \frac{1}{n} (\mathcal{M}(\mathbf{x}^{(i)}) - \bar{Y})^2$$

This error estimator leads to overfitting.
Overfitting – Illustration of the Runge effect

- When using a polynomial regression model that uses the same number of points as the degree of the polynomial, one gets an interpolating approximation.

- The empirical error is zero whereas the approximation gets worse and worse.
Error estimators

Leave-one-out cross validation

Principle

- In statistical learning theory, cross validation consists in splitting the experimental design $\mathcal{Y}$ in two parts, namely a training set (which is used to build the model) and a validation set.

- The leave-one-out technique consists in using each point of the experimental design as a single validation point for the meta-model built from the remaining $n - 1$ points.

- $n$ different meta-models are built and the error made on the remaining point is computed, then mean-square averaged.
Cross validation
Implementation

- For each $x^{(i)}$, a polynomial chaos expansion is built using the following experimental design: $\mathcal{X} \setminus x^{(i)} = \{x^{(j)} : j = 1, \ldots, n, j \neq i\}$, denoted by $\mathcal{M}^{PC\setminus i}(\cdot)$.

- The predicted residual is computed in point $x^{(i)}$:

$$\Delta_i = \mathcal{M}(x^{(i)}) - \mathcal{M}^{PC\setminus i}(x^{(i)})$$

- The PRESS coefficient (predicted residual sum of squares) is evaluated:

$$PRESS = \sum_{i=1}^{n} \Delta_i^2$$

- The leave-one-out error and related $Q^2$ error estimator are computed:

$$E_{LOO} = \frac{1}{n} \sum_{i=1}^{n} \Delta_i^2 \quad Q^2 = 1 - \frac{E_{LOO}}{\hat{\nu}[Y]}$$
In practice one does not need to explicitly derive the $n$ different models $\mathcal{M}_{PC\setminus i}(\cdot)$.

In contrast, a single least-square analysis is carried out. The predicted residual reads:

$$\Delta_i = \mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}_{PC\setminus i}(\mathbf{x}^{(i)}) = \frac{\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}_{PC}(\mathbf{x}^{(i)})}{1 - h_i}$$

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

$$\mathbf{A}_{ij} = \Psi_j(\mathbf{x}^{(i)})$$

Thus:

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

1. Sparse polynomial chaos scheme
2. Adaptive algorithms for sparse expansions
3. Sensitivity analysis
4. Application examples in sensitivity analysis
Index of sparsity

- Effect of the hyperbolic truncation scheme

\[ IS1 = \frac{\text{card } \mathcal{A}_{q,p}^M}{\text{card } \mathcal{A}_{M,p}} \]
Index of sparsity

- **Effect of the hyperbolic truncation scheme**

\[ IS_1 = \frac{\text{card } \mathcal{A}_q^{M,p}}{\text{card } \mathcal{A}^{M,p}} \]

- **Effect of the adaptive selection algorithm**

\[ IS_2 = \frac{\text{card } \mathcal{A}}{\text{card } \mathcal{A}_q^{M,p}} \]
How to get sparse expansions?

- Finding the significant coefficients in the PC expansion is a variable selection problem.
- It can be addressed by penalized regression techniques: ridge regression ($L_2$ penalty term), LASSO ($L_1$ penalty term)
- The Least Angle Regression (LAR) algorithm is an efficient approach:
  - A set of candidate basis functions $\mathcal{A}$ is pre-selected, e.g. using the hyperbolic truncation scheme
  - A family of sparse models are built, containing resp. 1, 2, $\ldots$, $|\mathcal{A}|$ terms
  - Among these models the best one is retained by applying cross validation techniques
Consider a 3-dimensional vector
The algorithm is initialized with $\hat{f}^0 = 0$. The residual is $R = y$.

The most correlated regressor is $X_1$. 

Efron et al. (2004)

Guigue et al. (2006)
A move in the direction $X_1$ is carried out so that the residual $y - \gamma_1 X_1$ becomes equicorrelated with $X_1$ and $X_2$.

The 1-term sparse approximation of $y$ is $\hat{f}^1$. 

Efron et al. (2004)
Guigue et al. (2006)
• A move is jointly made in the directions $X_1, X_2$ until the residual becomes equicorrelated with a new regressor
• This gives the 2-term sparse approximation
Least angle regression

Path of solutions

- A path of solutions is obtained containing $1, 2, \ldots, \min(n, \text{Card } \mathcal{A})$ terms.
- A $Q^2$-estimator of the accuracy of each solution is evaluated by cross validation and the best result is kept.
Basis-and-design adaptive LAR

**Initialization**
- Choose a norm $q$, $0 < q \leq 1$
- Select an initial design $\mathcal{X}$
- Store the model evaluations in $\mathcal{Y}$

**Selection of an optimal PC basis $\mathcal{A}^*$**
- For $p = 1, \ldots, p_{\text{max}}$:
  - Apply LAR to the candidate basis $\mathcal{A}$ which contains all those terms with $q$-norm $\leq p$
  - Let $\mathcal{A}^{(p)}$ be the optimal basis obtained by LAR and $\epsilon_{\text{LOO}}^*$ the corresponding error estimate (corrected leave-one-out estimate)
  - Store $\epsilon_{\text{LOO}, \text{min}}^* \equiv \min(\epsilon_{\text{LOO}}^*)$ and the associated basis $\mathcal{A}_{\text{min}}$

**STOP** if $\epsilon_{\text{LOO}, \text{min}}^*$ is less than a target error $\epsilon_{\text{tgt}}$

**Enrich the design $\mathcal{X}$** if $\epsilon_{\text{LOO}}^*$ increases twice in a row (overfitting). Restart the procedure from degree $p_{\text{min}}$ of basis $\mathcal{A}_{\text{min}}$

**Compute the coefficients** associated with $\mathcal{A}_{\text{min}}$ by least-square regression
Outline

1. Sparse polynomial chaos scheme
2. Adaptive algorithms for sparse expansions
3. Sensitivity analysis
   - Sobol’ indices
   - Case of dependent inputs: ANCOVA
4. Application examples in sensitivity analysis
Sensitivity analysis

Sobol’ decomposition

- Sensitivity analysis aims at quantifying what are the input parameters (or combinations thereof) that influence the most the response variability.
- Global sensitivity analysis relies on so-called variance decomposition techniques.

Consider a model \( M : \mathbf{x} \in [0, 1]^M \to M(\mathbf{x}) \in \mathbb{R} \). The Hoeffding-Sobol’ decomposition reads:

\[
M(\mathbf{x}) = M_0 + \sum_{i=1}^{M} M_i(x_i) + \sum_{1 \leq i < j \leq M} M_{ij}(x_i, x_j) + \cdots + M_{12\ldots M}(\mathbf{x})
\]

where:
- \( M_0 \) is the mean value of the function
- \( M_i(x_i) \) are univariate functions
- \( M_{ij}(x_i, x_j) \) are bivariate functions
- etc.
The functional decomposition is unique if the orthogonality of the terms (with respect to the uniform measure) is imposed \([0, 1]^M\), i.e. 
\[\{i_1, \ldots, i_s\} \neq \{j_1, \ldots, j_t\} \]:

\[
\int_{[0,1]^M} M_{i_1 \ldots i_s}(x_{i_1}, \ldots, x_{i_s}) M_{j_1 \ldots j_t}(x_{j_1}, \ldots, x_{j_t}) \, dx = 0
\]

A construction definition of the terms is obtained by recurrence:

\[
M_i(x_i) = \int_{[0,1]^{M-1}} M(x) \, dx_{\sim i} - M_0
\]

\[
M_{ij}(x_i, x_j) = \int_{[0,1]^{M-2}} M(x) \, dx_{\sim \{ij\}} - M_i(x_i) - M_j(x_j) - M_0
\]

where \(\int_{[0,1]^{M-1}} (.) \, dx_{\sim i}\) denoted the integration over all variables except for the \(i\)-th one.
Sobol’ indices

Variance decomposition

- Assume $X_i \sim \mathcal{U}(0, 1)$, $i = 1, \ldots, M$ (possibly after some isoprobabilistic transform)

- Due to the orthogonality of the decomposition:

\[
D \equiv \text{Var}[\mathcal{M}(X)] = \mathbb{E}[(\mathcal{M}(X) - \mathcal{M}_0)^2] = \mathbb{E}\left[\left(\sum_{\{i_1, \ldots, i_s\} \subset \{1, \ldots, M\}} \mathcal{M}_{i_1 \ldots i_s}(X_{i_1}, \ldots, X_{i_s})\right)^2\right]
\]

\[
= \sum_{\{i_1, \ldots, i_s\} \subset \{1, \ldots, M\}} \mathbb{E}\left[\mathcal{M}_{i_1 \ldots i_s}^2(X_{i_1}, \ldots, X_{i_s})\right]
\]
Sobol’ indices

Partial variance

- Consider:

\[ D_{i_1 \ldots i_s} = \int_{[0,1]^s} M_{i_1 \ldots i_s}^2 (x_{i_1}, \ldots, x_{i_s}) \, dx_{i_1} \ldots dx_{i_s} \]

- Then:

\[ D \equiv \text{Var}[Y] = \sum_{i=1}^{M} D_i + \sum_{1 \leq i < j \leq M} D_{ij} + \ldots + D_{12\ldots M} \]

- The Sobol’ indices are obtained by normalization:

\[ S_{i_1 \ldots i_s} = \frac{D_{i_1 \ldots i_s}}{D} \]

They represent the fraction of the total variance \( \text{Var}[Y] \) that can be attributed to each input variable \( i \) (\( S_i \)) or combinations of variables \( \{i_1 \ldots i_s\} \).
First order and total Sobol’ indices

First order Sobol’ indices

\[ S_i = \frac{D_i}{D} \quad D_i = \text{Var}_X \left[ M_i(X) \right] = \text{Var}_X \left[ \mathbb{E} [M(X)|X_i = x_i] \right] \]

They quantify the (additive) effect of each input parameter separately, i.e. the reduction of variance to expect if \(X_i\) is set equal to \(x_i\).

Total Sobol’ indices

\[ S_i^T \overset{\text{def}}{=} \sum_{i \subset \{i_1 \ldots i_s\}} S_{i_1 \ldots i_s} \]

They quantify the total effect of \(X_i\) including the first order effect and the interactions with the other variables.
Consider \( Y = \mathcal{M}(X) \) where \( X \sim f_X \) with independent components:

\[
Y = \sum_{\alpha \in \mathbb{N}^M} y_{\alpha} \Psi_{\alpha}(X)
\]

Due to the orthogonality properties of the polynomial chaos basis, one gets:

\[
\mathbb{E}[\Psi_{\alpha}(X)] = 0 \quad \mathbb{E}[\Psi_{\alpha}(X)\Psi_{\beta}(X)] = \delta_{\alpha\beta}
\]

Thus the mean and variance:

\[
\mathcal{M}_0 = \mathbb{E}[\mathcal{M}(X)] = y_0
\]

\[
D = \text{Var}[\mathcal{M}(X)] = \sum_{\alpha \in \mathbb{N}^M, \alpha \neq 0} y_{\alpha}^2
\]
Interaction sets

Let $\mathcal{A}_u$ be the set of multi-indices depending \textit{exactly} on the subset of variables $u = \{i_1, \ldots, i_s\}$:

$$\mathcal{A}_u = \{ \alpha \in \mathbb{N}^M : k \in u \iff \alpha_k \neq 0 \}$$

$$\bigcup_{u \subseteq \{1, \ldots, M\}} \mathcal{A}_u = \mathbb{N}^M$$

Sobol’ decomposition

By unicity of the Sobol’ decomposition one gets ($x_u \overset{\text{def}}{=} \{x_{i_1}, \ldots, x_{i_s}\}$):

$$\mathcal{M}(x) = \mathcal{M}_0 + \sum_{u \subseteq \{1, \ldots, M\}} \mathcal{M}_u(x_u)$$

where:

$$\mathcal{M}_u(x_u) \overset{\text{def}}{=} \sum_{\alpha \in \mathcal{A}_u} y_{\alpha} \Psi_\alpha(x)$$
Sparse polynomial chaos scheme
Adaptive algorithms for sparse expansions
Sensitivity analysis
Application examples

Partial variances

The partial variances $D_u \overset{\text{def}}{=} D_{i_1 \ldots i_s} = \text{Var}[\mathcal{M}_u(X)]$ are obtained by summing up the square of selected PC coefficients.

First order contribution

$$D_i = \sum_{\alpha \in A_i} y_\alpha^2 \quad A_i = \{ \alpha \in \mathbb{N}^M : \alpha_i > 0, \alpha_j \neq i = 0 \}$$

Higher order contribution

$$D_u = \sum_{\alpha \in A_u} y_\alpha^2 \quad A_{i_1 \ldots i_s} = \{ \alpha \in \mathbb{N}^M : k \in u \Leftrightarrow \alpha_k > 0 \}$$

- The Sobol’ indices come after normalization:

$$S_u = \frac{D_u}{D}$$
Sobol’ decomposition from PC expansions

First order indices

\[ S_i = \sum_{\alpha \in A_i} \frac{y_\alpha^2}{D} \quad A_i = \{ \alpha \in \mathbb{N}^M : \alpha_i > 0, \alpha_j \neq i = 0 \} \]

Higher order indices

\[ S_{i_1, \ldots, i_s} = \sum_{\alpha \in A_{i_1, \ldots, i_s}} \frac{y_\alpha^2}{D} \quad A_{i_1, \ldots, i_s} = \{ \alpha \in \mathbb{N}^M : k \in \{i_1, \ldots, i_s\} \Leftrightarrow \alpha_j \neq 0 \} \]

Total indices

\[ S_i^T = \sum_{\alpha \in A_i^T} \frac{y_\alpha^2}{D} \quad A_i^T = \{ \alpha \in \mathbb{N}^M : \alpha_i > 0 \} \]
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

\[ p = 3, \text{ i.e. } P = 10 \text{ terms} \]

\[ j \quad \alpha \quad \psi_{\alpha} \equiv \psi_j \]

<table>
<thead>
<tr>
<th>( j )</th>
<th>( \alpha )</th>
<th>( \psi_{\alpha} \equiv \psi_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>[0, 0]</td>
<td>( \psi_0 = 1 )</td>
</tr>
<tr>
<td>1</td>
<td>[1, 0]</td>
<td>( \psi_1 = \xi_1 )</td>
</tr>
<tr>
<td>2</td>
<td>[0, 1]</td>
<td>( \psi_2 = \xi_2 )</td>
</tr>
<tr>
<td>3</td>
<td>[2, 0]</td>
<td>( \psi_3 = (\xi_1^2 - 1)/\sqrt{2} )</td>
</tr>
<tr>
<td>4</td>
<td>[1, 1]</td>
<td>( \psi_4 = \xi_1 \xi_2 )</td>
</tr>
<tr>
<td>5</td>
<td>[0, 2]</td>
<td>( \psi_5 = (\xi_2^2 - 1)/\sqrt{2} )</td>
</tr>
<tr>
<td>6</td>
<td>[3, 0]</td>
<td>( \psi_6 = (\xi_3^3 - 3\xi_1)/\sqrt{6} )</td>
</tr>
<tr>
<td>7</td>
<td>[2, 1]</td>
<td>( \psi_7 = (\xi_2^2 - 1)\xi_2/\sqrt{2} )</td>
</tr>
<tr>
<td>8</td>
<td>[1, 2]</td>
<td>( \psi_8 = (\xi_2^2 - 1)\xi_1/\sqrt{2} )</td>
</tr>
<tr>
<td>9</td>
<td>[0, 3]</td>
<td>( \psi_9 = (\xi_2^3 - 3\xi_2)/\sqrt{6} )</td>
</tr>
</tbody>
</table>

Variance

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

Sobol’ indices

\[ S_1 = \frac{a_1^2 + a_3^2 + a_6^2}{D} \]

\[ S_2 = \frac{a_2^2 + a_5^2 + a_9^2}{D} \]

\[ S_{12} = \frac{a_4^2 + a_7^2 + a_8^2}{D} \]
Outline

1. Sparse polynomial chaos scheme
2. Adaptive algorithms for sparse expansions
3. Sensitivity analysis
   - Sobol’ indices
   - Case of dependent inputs: ANCOVA
4. Application examples in sensitivity analysis
Consider $Y = \mathcal{M}(X)$ where $X$ has a joint PDF $f_X$ (i.e. with dependent components)

Assume a functional decomposition exists (e.g. the High Dimensional Model Representation (HDMR), see Rabitz et al.):

$$\mathcal{M}(x) = m_0 + \sum_{i=1}^{M} m_i(x_i) + \sum_{1 \leq i < j \leq M} m_{ij}(x_i, x_j) + \ldots$$

The various terms in the expansion are not necessarily orthogonal anymore with respect to $\mathbb{P}(dx) = f_X(x) \, dx$:

$$\mathbb{E}_X [m_u(X_u) \, m_v(X_v)] \neq 0$$

The covariance decomposition is an extension of the Sobol’ decomposition accounting from the cross terms
Covariance decomposition

\[
\text{Var} [m(X)] = \text{Cov} \left[ m_0 + \sum_{u \subseteq \{1, \ldots, M\}} m_u(X_u), m(X) \right]
= \sum_{u \subseteq \{1, \ldots, M\}} \text{Cov} [m_u(X_u), m(X)]
\]

**ANCOVA sensitivity indices**

\[
S_u^{(cov)} = \frac{\text{Cov} [m_u(X_u), m(X)]}{\text{Var} [m(X)]} = \frac{\text{Var} [m_u(X_u)] + \sum_{v \neq u} \text{Cov} [m_u(X_u); m_v(X_v)]}{\text{Var} [m(X)]}
\]

**Structural (uncorrelated):**

\[
S_u^{(U)} = \frac{\text{Var} [m_u(X_u)]}{\text{Var} [m(X)]}
\]

**Correlated:**

\[
S_u^{(C)} = \frac{\sum_{v \neq u} \text{Cov} [m_u(X_u); m_v(X_v)]}{\text{Var} [m(X)]}
\]
Estimation from PC expansions

- A functional decomposition \( m(x) \overset{\text{def}}{=} M_A(x) \) is obtained by computing a truncated PC expansion of \( M(Z) \) where \( Z \) is made of independent variables of marginal PDF \( f_{X_i} \)

\[
f_Z(z) = \prod_{i=1}^{M} f_{X_i}(z_i)
\]

- The moments of \( Y = M(X) \) are computed by Monte Carlo simulation using a \( n \)-sample drawn according to \( f_X \) (i.e. with correlation):

\[
\overline{y_A} = \frac{1}{n} \sum_{i=1}^{n} M_A(x_i)
\]

\[
\text{Var}[Y_A] = \frac{1}{n-1} \sum_{i=1}^{n} (M_A(x_i) - \overline{y_A})^2
\]

\( \overline{y_A} \neq y_0 \) since \( \mathbb{E}_X[\Psi_\alpha(X)] \neq \mathbb{E}_Z[\Psi_\alpha(Z)] \)
Monte Carlo estimators of ANCOVA indices

**ANCOVA index**

\[ \bar{y}_u \overset{\text{def}}{=} \hat{E} [M_u(x_u)] = \frac{1}{n} \sum_{i=1}^{n} M_u(x_{u,i}) \]

\[ S_u^{(cov)} = \frac{\sum_{i=1}^{n} (M_A(x_i) - \bar{y}_A) (M_u(x_{u,i}) - \bar{y}_u)} {\sum_{i=1}^{n} (M_A(x_i) - \bar{y}_A)^2} \]

**Structural contribution**

\[ S_u^{(U)} = \frac{\sum_{i=1}^{n} (M_u(x_{u,i}) - \bar{y}_u)^2} {\sum_{i=1}^{n} (M_A(x_i) - \bar{y}_A)^2} \]

**Correlated contribution**

\[ S_u^{(C)} = S_u^{(cov)} - S_u^{(U)} \]
Outline

1. Sparse polynomial chaos scheme
2. Adaptive algorithms for sparse expansions
3. Sensitivity analysis
4. Application examples in sensitivity analysis
   - Ishigami function
   - Morris function
   - Bending beam
Ishigami function

Definition

\[ Y = \sin X_1 + a \sin^2 X_2 + b X_3^4 \sin X_1 \quad a = 7, \ b = 0.1 \]

where \( X_i \sim \mathcal{U}[-\pi, \pi] \) are independent uniform random variables

Analytical solution

\[ D = \frac{a^2}{8} + \frac{b \pi^4}{5} + \frac{b^2 \pi^8}{18} + \frac{1}{2} \]

\[ D_1 = \frac{b \pi^4}{5} + \frac{b^2 \pi^8}{50} + \frac{1}{2}, \quad D_2 = \frac{a^2}{8}, \quad D_3 = 0 \]

\[ D_{12} = D_{23} = 0, \quad D_{13} = \frac{8 b^2 \pi^8}{225}, \quad D_{123} = 0 \]
First order Sobol’ indices

<table>
<thead>
<tr>
<th>Index</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>0.3138</td>
</tr>
<tr>
<td>$S_2$</td>
<td>0.4424</td>
</tr>
<tr>
<td>$S_3$</td>
<td>0</td>
</tr>
<tr>
<td>$S_{12}$</td>
<td>0</td>
</tr>
<tr>
<td>$S_{13}$</td>
<td>0.2436</td>
</tr>
<tr>
<td>$S_{23}$</td>
<td>0</td>
</tr>
</tbody>
</table>
First order Sobol’ indices

\[ |S_i - S_i^{(ref)}| \]

Sample size $N$

\[ S_1 \ (MCS) \]
\[ S_1 \ (PC) \]
\[ S_2 \ (MCS) \]
\[ S_2 \ (PC) \]
\[ S_{13} \ (MCS) \]
\[ S_{13} \ (PC) \]
Total Sobol’ indices

\[ S_i^T \]

Sample size \( N \):

- \( S_1^T \) (MCS)
- \( S_1^T \) (PC)
- \( S_2^T \) (MCS)
- \( S_2^T \) (PC)
- \( S_3^T \) (MCS)
- \( S_3^T \) (PC)
Total Sobol’ indices – small design and replication

![Graph 1](image1)

![Graph 2](image2)

![Graph 3](image3)
Morris function
Definition

\[ Y = \beta_0 + \sum_{i=1}^{20} \beta_i w_i + \sum_{i<j}^{20} \beta_{ij} w_i w_j + \sum_{i<j<l}^{20} \beta_{ijl} w_i w_j w_l + \sum_{i<j<l<s}^{20} \beta_{ijls} w_i w_j w_l w_s \]

where:

\[ w_i = \begin{cases} 
  2 \left(1.1 \frac{X_i}{X_i + 0.1} - 0.5\right) & \text{if } i = 3, 5, 7 \\
  2(X_i - 0.5) & \text{otherwise}
\end{cases} \]

\[ X_i \sim \mathcal{U}(0, 1) \]

and:

\[ \begin{cases} 
  \beta_i = 20 & \text{for } i = 1, \ldots, 10 \\
  \beta_{ij} = -15 & \text{for } i = 1, \ldots, 6 \\
  \beta_{ijl} = -10 & \text{for } i = 1, \ldots, 5 \\
  \beta_{ijls} = 5 & \text{for } i = 1, \ldots, 4 \\
  \beta_i = (-1)^i & \text{otherwise} \\
  \beta_{ij} = (-1)^{i+j} & \text{otherwise} \\
  \beta_{ijl} = 0 & \text{otherwise} \\
  \beta_{ijls} = 0 & \text{otherwise}
\end{cases} \]
Morris function

Sensitivity results

- Reference: $N = 440,000$ Monte Carlo simulations + bootstrap
- Adaptive sparse PC: $Q^2_{tgt} = 0.9$ (resp. 0.99)
Morris function

Sparsity of the model

<table>
<thead>
<tr>
<th>Sensitivity indices</th>
<th>Sparse PCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Q_{\text{final}}^2 = 0.9$</td>
</tr>
<tr>
<td>$S_4$</td>
<td>0.26</td>
</tr>
<tr>
<td>$S_1$</td>
<td>0.25</td>
</tr>
<tr>
<td>$S_2$</td>
<td>0.24</td>
</tr>
<tr>
<td>$S_6$</td>
<td>0.16</td>
</tr>
<tr>
<td>$S_3$</td>
<td>0.10</td>
</tr>
<tr>
<td>$S_5$</td>
<td>0.08</td>
</tr>
<tr>
<td>$S_8$</td>
<td>0.11</td>
</tr>
<tr>
<td>$S_{10}$</td>
<td>0.10</td>
</tr>
<tr>
<td>$S_6$</td>
<td>0.10</td>
</tr>
<tr>
<td>$S_7$</td>
<td>0.08</td>
</tr>
<tr>
<td># model evaluations</td>
<td>250</td>
</tr>
<tr>
<td>PC degree</td>
<td>3</td>
</tr>
<tr>
<td>Index of sparsity $IS$</td>
<td>$250/1771 \approx 0.14$</td>
</tr>
</tbody>
</table>

- **Full chaos:**
  \[ \text{card } \mathcal{A}^{20,11} = \binom{20+11}{11} = 84,672,315 \]

- **Hyperbolic truncature:**
  \[ \text{card } \mathcal{A}^{20,11}_{0.4} = 4,234 \]

LAR: 339 terms ($IS_2 = 8\%$)
Outline

1. Sparse polynomial chaos scheme
2. Adaptive algorithms for sparse expansions
3. Sensitivity analysis
4. Application examples in sensitivity analysis
   - Ishigami function
   - Morris function
   - Bending beam
Elastic bending beam

Problem statement

Spatially varying Young’s modulus modelled by a stationary lognormal random field:
\[ \mu_E = 210,000 \text{ MPa, } CV_E = 20\% \]

- Input random field:
  \[ E(x, \omega) = \exp [\lambda_E + \zeta_E g(x, \omega)] \]
  \[ g(x, \omega) \sim \mathcal{N}(0, 1) \]
  \[ \text{Cov} [g(x) g(x')] = e^{-|x' - x| / \ell} \quad (\ell = 0.5 \text{ m } = L/6) \]

- Quantity of interest: maximal deflection at midspan
Karhunen-Loève expansion of the random field

- Analytical solution of the Fredholm integral equation
- 100 terms retained for a discretization accuracy $\sim 1\%$
## Beam deflection

<table>
<thead>
<tr>
<th></th>
<th>Reference</th>
<th>LAR - $\varepsilon_{tgt} = 0.01$</th>
<th>LAR - $\varepsilon_{tgt} = 0.001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (mm)</td>
<td>2.83</td>
<td>2.81</td>
<td>2.83</td>
</tr>
<tr>
<td>Standard Deviation (mm)</td>
<td>0.37</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>Skewness</td>
<td>[0.38 ; 0.52] †</td>
<td>0.30</td>
<td>0.41</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>[3.12 ; 3.70] †</td>
<td>3.09</td>
<td>3.30</td>
</tr>
<tr>
<td>Number of FE runs</td>
<td>10,000</td>
<td>200</td>
<td>1,200</td>
</tr>
<tr>
<td>Error estimate</td>
<td>$10^{-3}$</td>
<td>$4 \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>PC degree</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Number of PC terms</td>
<td>10</td>
<td>246</td>
<td></td>
</tr>
<tr>
<td>Index of sparsity $IS_1$</td>
<td>$2 \times 10^{-3}$</td>
<td>$3 \times 10^{-6}$</td>
<td></td>
</tr>
<tr>
<td>Index of sparsity $IS_2$</td>
<td>7%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>

- PC expansion of max. degree $p = 6$ in 100 dimensions:
  $$P = 1,705, 904, 746 \text{ terms}$$

- Hyperbolic truncature: 5,118 terms

- LAR: 246 terms
Only the modes #1 and #3 impact the maximal deflection (modes #2 #4 are antisymmetric).

Another analysis using only 3 terms in the KL expansion (and $N = 40$ finite element runs) provide the same second moments within 1% accuracy ($L_2$-error of $2.10^{-4}$)
Conclusions

- Polynomial chaos expansions suffer from the curse of dimensionality when classical Galerkin/ stochastic collocation techniques are used.

- Hyperbolic truncation schemes allow one to select a priori approximation spaces of lower dimensionality that are consistent with the sparsity-of-effect principle.

- Sparse expansions may be obtained by casting the problem of computing the PC coefficients in terms of regression analysis and variable selection (Least Angle Regression).

- Although stochastic problems (e.g. involving random fields) may appear as large dimensional, the effective dimension for quantities of interest is usually much smaller.
Questions ?

Thank you very much for your attention!

http://www.rsuq.ethz.ch

UQLab ...

... The Uncertainty Quantification Toolbox in Matlab