# Surrogate models for uncertain dynamical systems: applications to earthquake engineering

Bruno Sudret

**Chair of Risk, Safety and Uncertainty Quantification**

**ETH Zurich**

**Symposium on Uncertainty Quantification in Computational Geosciences**

BRGM (Orléans, France) – January 16th, 2018

# 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

# 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

# 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

# Global framework for uncertainty quantification

**Step B**
Quantification of sources of uncertainty

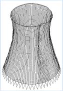
**Step A**
Model(s) of the system
Assessment criteria

**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)*

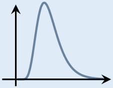
# Global framework for uncertainty quantification

**Step B**
Quantification of sources of uncertainty

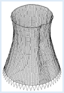
**Step A**
Model(s) of the system
Assessment criteria

**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)*

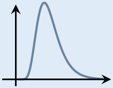
# Global framework for uncertainty quantification

**Step B**
Quantification of sources of uncertainty

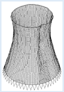
**Step A**
Model(s) of the system
Assessment criteria

**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)*

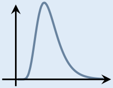
# Global framework for uncertainty quantification

**Step B**
Quantification of sources of uncertainty

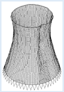
**Step A**
Model(s) of the system
Assessment criteria

**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)*

# Surrogate models for uncertainty quantification

A surrogate model $\tilde{\mathcal{M}}$ is an approximation of the original computational model $\mathcal{M}$ with the following features:

- It is built from a limited set of runs of the original model $\mathcal{M}$ called the experimental design $\mathcal{X} = \left\{ \boldsymbol{x}^{(i)},\, i = 1, \ldots, n \right\}$
- It assumes some regularity of the model $\mathcal{M}$ and some general functional shape

# Surrogate models for uncertainty quantification

A surrogate model $\tilde{\mathcal{M}}$ is an approximation of the original computational model $\mathcal{M}$ with the following features:

- It is built from a limited set of runs of the original model $\mathcal{M}$ called the experimental design $\mathcal{X} = \left\{ \boldsymbol{x}^{(i)},\, i = 1, \, \dots \,, n \right\}$

- It assumes some regularity of the model $\mathcal{M}$ and some general functional shape

# Surrogate models for uncertainty quantification

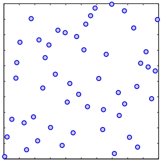
A surrogate model $\tilde{\mathcal{M}}$ is an approximation of the original computational model $\mathcal{M}$ with the following features:

- It is built from a limited set of runs of the original model $\mathcal{M}$ called the experimental design $\mathcal{X} = \left\{ \boldsymbol{x}^{(i)},\, i = 1, \, \dots \,, n \right\}$
- It assumes some regularity of the model $\mathcal{M}$ and some general functional shape

| **Name** | **Shape** | **Parameters** |
|---|---|---|
| Polynomial chaos expansions | $\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$ | $\boldsymbol{a}_{\boldsymbol{\alpha}}$ |
| Low-rank tensor approximations | $\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{l=1}^{R} b_l \left( \prod_{i=1}^{M} v_l^{(i)}(x_i) \right)$ | $b_l, \; z_{k,l}^{(i)}$ |
| Kriging (a.k.a Gaussian processes) | $\tilde{\mathcal{M}}(\boldsymbol{x}) = \boldsymbol{\beta}^{\mathsf{T}} \cdot \boldsymbol{f}(\boldsymbol{x}) + Z(\boldsymbol{x}, \omega)$ | $\boldsymbol{\beta} \, , \, \sigma_Z^2 \, , \, \boldsymbol{\theta}$ |
| Support vector machines | $\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{i=1}^{n} a_i \, K(\boldsymbol{x}_i, \boldsymbol{x}) + b$ | $\boldsymbol{a} \, , \, b$ |

# Ingredients for building a surrogate model

- Select an experimental design $\mathcal{X}$ that covers at best the domain of input parameters: Latin hypercube sampling (LHS), low-discrepancy sequences



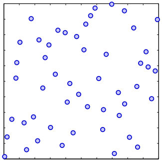
- Run the computational model $\mathcal{M}$ onto $\mathcal{X}$ exactly as in Monte Carlo simulation

- Smartly post-process the data $\{\mathcal{X}, \mathcal{M}(\mathcal{X})\}$ through a learning algorithm

| Name | Learning method |
|---|---|
| Polynomial chaos expansions | sparse grid integration, least-squares, compressive sensing |
| Low-rank tensor approximations | alternate least squares |
| Kriging | maximum likelihood, Bayesian inference |
| Support vector machines | quadratic programming |

# Ingredients for building a surrogate model

- Select an experimental design $\mathcal{X}$ that covers at best the domain of input parameters: Latin hypercube sampling (LHS), low-discrepancy sequences



- Run the computational model $\mathcal{M}$ onto $\mathcal{X}$ exactly as in Monte Carlo simulation

- Smartly post-process the data $\{\mathcal{X}\,,\,\mathcal{M}(\mathcal{X})\}$ through a learning algorithm

| Name | Learning method |
|---|---|
| Polynomial chaos expansions | sparse grid integration, least-squares, compressive sensing |
| Low-rank tensor approximations | alternate least squares |
| Kriging | maximum likelihood, Bayesian inference |
| Support vector machines | quadratic programming |

# Advantages of surrogate models

## Usage

$$\underset{\text{hours per run}}{\mathcal{M}(\boldsymbol{x})} \quad \approx \quad \underset{\text{seconds for } 10^6 \text{ runs}}{\tilde{\mathcal{M}}(\boldsymbol{x})}$$

## Advantages

- Non-intrusive methods: based on runs of the computational model, exactly as in Monte Carlo simulation
- Construction suited to high performance computing: "embarrassingly parallel"

# Advantages of surrogate models

## Usage

$$\mathcal{M}(\boldsymbol{x}) \quad \approx \quad \tilde{\mathcal{M}}(\boldsymbol{x})$$

hours per run — seconds for $10^6$ runs

## Advantages

- Non-intrusive methods: based on runs of the computational model, exactly as in Monte Carlo simulation
- Construction suited to high performance computing: "embarrassingly parallel"

## Challenges

- Need for rigorous validation
- Communication: advanced mathematical background

# Advantages of surrogate models

## Usage

$$\mathcal{M}(\boldsymbol{x}) \quad \approx \quad \tilde{\mathcal{M}}(\boldsymbol{x})$$

hours per run — seconds for $10^6$ runs

## Advantages

- Non-intrusive methods: based on runs of the computational model, exactly as in Monte Carlo simulation
- Construction suited to high performance computing: "embarrassingly parallel"

## Challenges

- Need for rigorous validation
- Communication: advanced mathematical background

Efficiency: 2-3 orders of magnitude less runs compared to Monte Carlo

# Outline

# Polynomial chaos expansions in a nutshell

Ghanem & Spanos (1991); Xiu & Karniadakis (2002); Soize & Ghanem (2004); Lemaître & Knio (2010)

- 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

# Computing the coefficients by least-square minimization

Isukapalli (1999); Berveiller, Sudret & Lemaire (2006)

## Principle

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

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

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

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

## Least-square minimization

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

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

# Computing the coefficients by least-square minimization

Isukapalli (1999); Berveiller, Sudret & Lemaire (2006)

## Principle

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

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

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

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

## Least-square minimization

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

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

# Discrete (ordinary) least-square minimization

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

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

## Procedure

- Select a truncation scheme, *e.g.* $\mathcal{A}^{M,p} = \left\{\boldsymbol{\alpha}\in\mathbb{N}^M \;:\; |\boldsymbol{\alpha}|_1 \le p\right\}$
- Select an experimental design and evaluate the model response

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

- Compute the experimental matrix

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

- Solve the resulting linear system

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

# Discrete (ordinary) least-square minimization

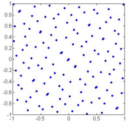
An estimate of the mean square error (sample average) is minimized:

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

## Procedure

- Select a truncation scheme, *e.g.* $\mathcal{A}^{M,p} = \left\{\boldsymbol{\alpha}\in\mathbb{N}^M \;:\; |\boldsymbol{\alpha}|_1 \leq p\right\}$
- Select an experimental design and evaluate the model response

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




- Compute the experimental matrix

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

- Solve the resulting linear system

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

# Discrete (ordinary) least-square minimization

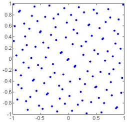
An estimate of the mean square error (sample average) is minimized:

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

## Procedure

- Select a truncation scheme, *e.g.* $\mathcal{A}^{M,p} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M \ : \ |\boldsymbol{\alpha}|_1 \le p \right\}$
- Select an experimental design and evaluate the model response

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



- Compute the experimental matrix

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

- Solve the resulting linear system

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

# Discrete (ordinary) least-square minimization

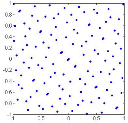
An estimate of the mean square error (sample average) is minimized:

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

## Procedure

- Select a truncation scheme, *e.g.* $\mathcal{A}^{M,p} = \left\{\boldsymbol{\alpha} \in \mathbb{N}^M \;:\; |\boldsymbol{\alpha}|_1 \leq p\right\}$
- Select an experimental design and evaluate the model response

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



- Compute the experimental matrix

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

- Solve the resulting linear system

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

Simple is beautiful !

# Error estimators

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

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

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

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

# 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

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

# 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

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

## Leave-one-out cross validation

- From statistical learning theory, model validation shall be carried out using independent data

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

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

# Outline

# Why brute-force PCE fails in dynamics?

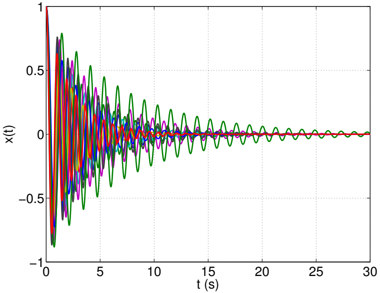
## Non-linear SDOF Duffing oscillator

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

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

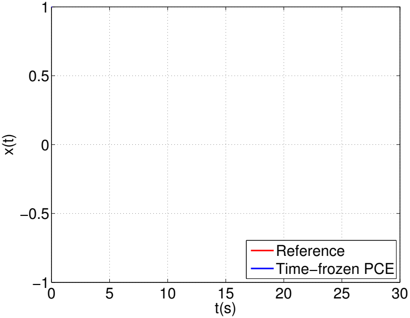
Input: 3 uniform random variables

| RV | Distribution | Values |
|---|---|---|
| $\zeta$ | Uniform | $\mathcal{U}[0.015,\, 0.045]$ |
| $\omega$ | Uniform | $\mathcal{U}[\pi,\, 3\,\pi]$ |
| $\varepsilon$ | Uniform | $\mathcal{U}[-0.25,\, -0.75]$ |

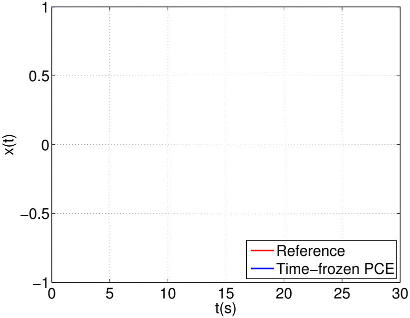



**Samples of trajectories**

# Time-frozen PCE



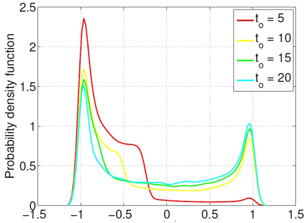

$(\zeta, \omega, \varepsilon) = (0.03, 8.92, -0.34)$

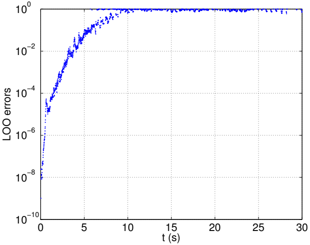



$(\zeta, \omega, \varepsilon) = (0.04, 3.18, -0.33))$

# Why time-frozen PCE does not work?

- The map $\boldsymbol{\xi} \mapsto \mathcal{M}(\boldsymbol{\xi}, t)$ becomes increasingly non linear with time
- The time-frozen distribution of the output at time $t_0$ becomes more complex (*e.g.* multimodal)
- Expansions of higher degree would be required to keep sufficient accuracy with time
- For a fixed experimental design, the LOO error blows up



# Outline

# Some literature

- Multi-elements PCEs: decomposition of the random space into non-overlapping sub-elements
  Wan & Karniadakis, 2005
- Constant phase interpolation: responses interpolated in the phase space
  Witteveen & Bijl, 2008
- Asynchronous time integration: intrusive transformed time variable introduced to reduce variability
  Le Maître et al., 2010
- Time-dependent PCEs: new random variables added on-the-fly
  Gerritsma et al., 2010
- PC flow map composition: long-term response obtained by composing intermediate PCE-based flow maps
  Luchtenburg et al., 2014
- PC-NARX: future state determined by current and past states
  Spiridonakos & Chatzi, 2015

# Some literature

- Multi-elements PCEs: decomposition of the random space into non-overlapping sub-elements — Wan & Karniadakis, 2005
- Constant phase interpolation: responses interpolated in the phase space — Witteveen & Bijl, 2008
- Asynchronous time integration: intrusive transformed time variable introduced to reduce variability — Le Maître et al., 2010
- Time-dependent PCEs: new random variables added on-the-fly — Gerritsma et al., 2010
- PC flow map composition: long-term response obtained by composing intermediate PCE-based flow maps — Luchtenburg et al., 2014
- PC-NARX: future state determined by current and past states — Spiridonakos & Chatzi, 2015

# Nonlinear AutoRegressive with eXogenous input model

## NARX model

Billings, 2013

Based on a time-dependent input excitation $x(t)$ and corresponding system response $y(t)$, the dynamics is captured through:

$$y(t) = \mathcal{F}\left(x(t), \ldots, x(t-n_x), y(t-1), \ldots, y(t-n_y)\right) + \varepsilon_t$$

where:

- $\boldsymbol{z}(t) = \left(x(t), \ldots, x(t-n_x), y(t-1), \ldots, y(t-n_y)\right)^{\mathsf{T}}$ is the vector of current and past values
- $n_x$ and $n_y$ denote the maximum input and output time lags
- $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2(t))$ is the residual error
- $\mathcal{F}(\cdot)$ is a functional of NARX terms, usually linear-in-parameters:

$$y(t) = \sum_{i=1}^{n_g} \vartheta_i \, g_i(\boldsymbol{z}(t)) + \varepsilon_t$$

# PC-NARX model

Spiridonakos *et al.* , 2015a,2015b

## Computational model with uncertainties

$$y(t, \boldsymbol{\xi}_x, \boldsymbol{\xi}_s) \stackrel{\text{def}}{=} \mathcal{M}(x(t, \boldsymbol{\xi}_x), \boldsymbol{\xi}_s)$$

- $\boldsymbol{\xi}_x$ : uncertainty in the input excitation
- $\boldsymbol{\xi}_s$ : uncertainty in the system

## PC-NARX expansion

$$y(t, \boldsymbol{\xi}) = \sum_{i=1}^{n_g} \vartheta_i(\boldsymbol{\xi})\, g_i(\boldsymbol{z}(t)) + \varepsilon_g(t, \boldsymbol{\xi}) \qquad \boldsymbol{\xi} = (\boldsymbol{\xi}_x, \boldsymbol{\xi}_s)$$

The NARX stochastic coefficients $\vartheta_i(\boldsymbol{\xi})$ are represented by PCEs:

$$\vartheta_i(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i} \vartheta_{i,\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

# PC-NARX model

Spiridonakos *et al.* , 2015a,2015b

## Computational model with uncertainties

$$y(t, \boldsymbol{\xi}_x, \boldsymbol{\xi}_s) \stackrel{\text{def}}{=} \mathcal{M}(x(t, \boldsymbol{\xi}_x), \boldsymbol{\xi}_s)$$

- $\boldsymbol{\xi}_x$ : uncertainty in the input excitation
- $\boldsymbol{\xi}_s$ : uncertainty in the system

## PC-NARX expansion

$$y(t, \boldsymbol{\xi}) = \sum_{i=1}^{n_g} \vartheta_i(\boldsymbol{\xi})\, g_i(\boldsymbol{z}(t)) + \varepsilon_g(t, \boldsymbol{\xi}) \qquad \boldsymbol{\xi} = (\boldsymbol{\xi}_x, \boldsymbol{\xi}_s)$$

The NARX stochastic coefficients $\vartheta_i(\boldsymbol{\xi})$ are represented by PCEs:

$$\vartheta_i(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i} \vartheta_{i,\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

# PC-NARX model

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

## Interpretation

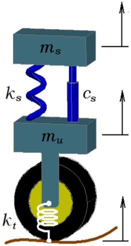
- PC-NARX is a NARX model in which each (random) coefficient is expanded as a PCE
- Compared to time-frozen PCE, a specific dynamics of the random coefficients is imposed
- Similar to flow map composition since the response at current instant is used to predict the response at future instants

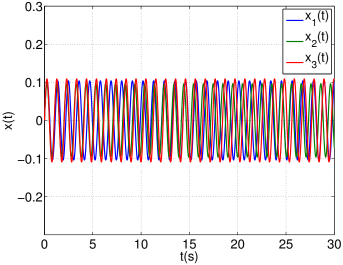
# Experimental design

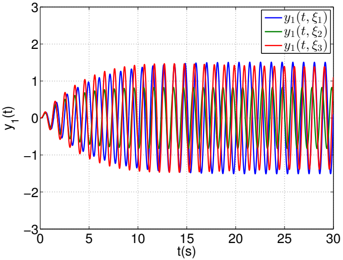
## Data

- $N$ realizations of the input excitation, cast as $\left(x_k[1], \ldots, x_k[T]\right)^{\mathsf{T}}, \ k = 1, \ldots, N$ ($T$ time instants)
- The corresponding system response computed by a simulator, cast as $\left(y_k[1], \ldots, y_k[T]\right)^{\mathsf{T}}$

## Example: quarter car model



# Deterministic NARX calibration

For a particular realization $\boldsymbol{\xi}_k$

- Select NARX model (candidate terms):

$$\boldsymbol{z}(t) = (x(t), \ldots, x(t-n_x), y(t-1), \ldots, y(t-n_y))^{\mathsf{T}}$$
$$\boldsymbol{\phi}(t) = \{g_i(\boldsymbol{z}(t)),\ i = 1, \ldots, n_g\}^{\mathsf{T}}$$

- Use least angle regression (LARS) to select the best explanatory subset of terms Efron *et al.* , 2004

- Compute the coefficients $\boldsymbol{\vartheta}_k$ by ordinary least-squares

Prediction error (of model $\#k$ on trajectory $l$)

$$\varepsilon_l^{\#k} = \frac{\sum_{t=1}^{T} (y(t,\boldsymbol{\xi}_l) - \hat{y}^{\#k}(t,\boldsymbol{\xi}_l))^2}{\sum_{t=1}^{T} (y(t,\boldsymbol{\xi}_l) - \bar{y}(t,\boldsymbol{\xi}_l))^2}$$

# Common NARX basis

## Premise

To expand the NARX coefficients onto a PC basis, it is necessary to have a common NARX model for all trajectories

## Procedure

- Select $K \leq N$ trajectories ("NARX learning set"), *e.g.* with the strongest non linear behaviour (peak displacement, velocities, etc.)
- Determine the sparse deterministic NARX models for realizations $k = 1, \ldots, K$, which leads to $P \leq K$ different possible models called $\#1, \ldots, \#P$
- Compute the NARX coefficients of the $N$ trajectories, for each model $\#p$, and evaluate an average error:

$$\varepsilon_p = \frac{1}{N} \sum_{k=1}^{N} \varepsilon_k^{\#p}$$

- Select the final best NARX model that minimizes $\varepsilon_p$

# PCE of the NARX coefficients

## PCE calibration

- Once a common NARX basis has been found, $N$ realizations of the NARX coefficients are available:

$$\mathcal{ED} = \{\vartheta_{i,k}, i = 1, \ldots, n_g;\ k = 1, \ldots, N\}$$

- $n_g$ different sparse PC expansions are built from this experimental design, using least-angle regression (LAR) Blatman & Sudret, 2011

$$\vartheta_i(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i} \vartheta_{i,\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

## PC-NARX prediction

- For a new realization of the input parameters $\boldsymbol{\xi}_0$, the NARX coefficients are first evaluated from PCEs
- Then they are plugged into the NARX model

# Bouc-Wen model

## Governing equations

Kafali & Grigoriu (2007), Spiridonakos & Chatzi (2015)

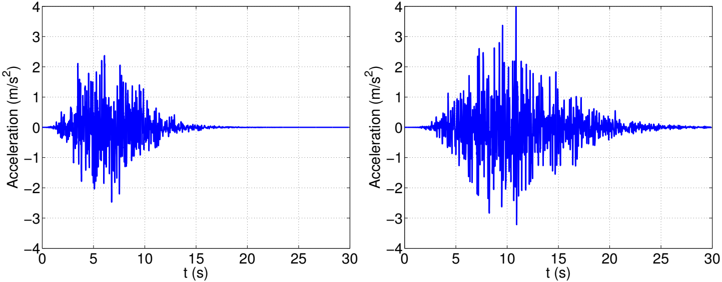
$$\ddot{y}(t) + 2\,\zeta\,\omega\,\dot{y}(t) + \omega^2(\rho\,y(t) + (1-\rho)\,z(t)) = -x(t),$$
$$\dot{z}(t) = \gamma\dot{y}(t) - \alpha\,|\dot{y}(t)|\,|z(t)|^{n-1}\,z(t) - \beta\,\dot{y}(t)\,|z(t)|^n\,,$$

## Excitation

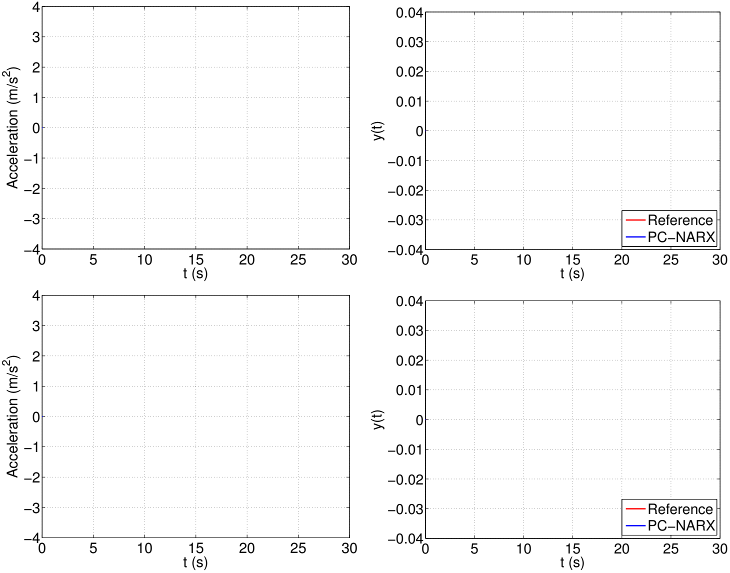
$x(t)$ is generated by a probabilistic ground motion model

Rezaeian & Der Kiureghian (2010)

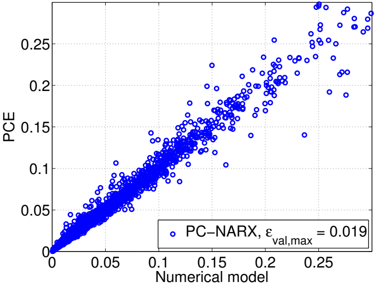
$$x(t) = q(t,\boldsymbol{\alpha}) \sum_{i=1}^{n} s_i\left(t, \boldsymbol{\lambda}(t_i)\right)\, U_i$$



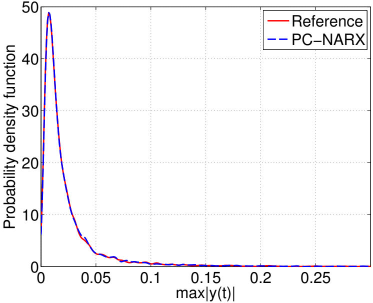
# Bouc-Wen model: prediction



# Bouc-Wen model: prediction




Maximal displacements




PDF of maximal displacements

# Outline

# Introduction to fragility curves

- Earthquake engineering aims at assessing the performance of structures and infrastructures w.r.t recorded or potential quakes
- Due to uncertainties in the localization, magnitude, structural behaviour and resistance, etc. probabilistic approaches are commonly used

## Fragility curves

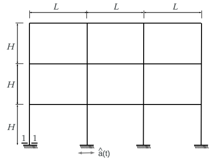
For a given performance criterion $g \leq g_{adm}$, the fragility curve represents the conditional probability of failure given an intensity measure $IM$:

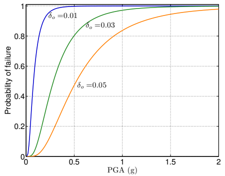
$$\text{Frag}(IM;\, g_{adm}) = \mathbb{P}\left(g \geq g_{adm} \,|\, IM\right)$$

**Example**

- $g = \max_k \max_{t_i \in [0,T]} |\delta^k_{t_i}|$ ($k$-th interstorey drift)
- $IM$: peak ground acceleration (PGA), pseudo-spectral acceleration (PSa), cumulative absolute velocity (CAV), etc.

# Fragility curves




## Classical approach

- Select a set of ground motions (recorded / synthetic)
- Compute the transient structural response (finite element analysis)
- Assume a parametric shape for the fragility curve, *e.g.* a lognormal shape:

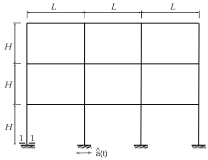
$$\text{Frag}(IM;\, \delta_o) = \mathbb{P}\left(\Delta \geq \delta_o \,|\, IM\right) = \Phi\left(\frac{\log IM - \alpha}{\beta}\right)$$

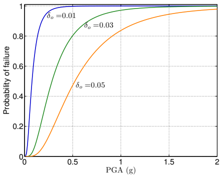
- Fit the parameters $(\alpha, \beta)$ form data

## Limitations

- Predefined shape of the curve
- Subject to epistemic uncertainties when the number of ground motions is small

# Fragility curves




## Classical approach

- Select a set of ground motions (recorded / synthetic)
- Compute the transient structural response (finite element analysis)
- Assume a parametric shape for the fragility curve, *e.g.* a lognormal shape:

$$\text{Frag}(IM;\, \delta_o) = \mathbb{P}\left(\Delta \geq \delta_o \,|\, IM\right) = \Phi\left(\frac{\log IM - \alpha}{\beta}\right)$$

- Fit the parameters $(\alpha, \beta)$ form data

## New proposal

- Use non parametric statistics for the fragility curves
- Use surrogate models of the transient analysis based on polynomial chaos expansions

# Parametric methods

## Linear regression (LR)

Ellingwood (2001)

- Probabilistic demand model:

$$\log\Delta = A\,\log IM + B + \zeta\,Z \qquad Z\sim\mathcal{N}(0,1)$$

- $A$ and $B$ determined by ordinary least squares estimation in a log-log scale
- Results in a lognormal-like fragility curve:

$$\widehat{\text{Frag}}(IM;\,\delta_o) = \mathbb{P}\left[\log\Delta \geq \log\delta_o\right] = 1 - \mathbb{P}\left[\log\Delta \leq \log\delta_o\right]$$
$$= \Phi\left(\frac{\log IM - \left(\log\delta_o - B\right)/A}{\zeta/A}\right).$$

## Maximum likelihood estimation (ML)

Shinozuka *et al.* (2000)

- Lognormal shape:

$$\widehat{\text{Frag}}(IM;\,\delta_o) = \Phi\left(\frac{\log IM - \log\alpha}{\beta}\right)$$

- Estimation of $\alpha$ and $\beta$ by maximum likelihood for each $\delta_o$:

$$\mathcal{L}\left(\alpha,\,\beta,\,\{IM_i\}_{i=1}^{N}\right) = \prod_{IM_i:\,\Delta_i\geq\delta_o}\left[\widehat{\text{Frag}}(IM_i;\,\delta_o)\right]\prod_{IM_i:\,\Delta_i<\delta_o}\left[1-\widehat{\text{Frag}}(IM_i;\,\delta_o)\right]$$

# Parametric methods

## Linear regression (LR)

Ellingwood (2001)

- Probabilistic demand model:

$$\log \Delta = A \, \log IM + B + \zeta \, Z \qquad Z \sim \mathcal{N}(0,1)$$

- $A$ and $B$ determined by ordinary least squares estimation in a log-log scale
- Results in a lognormal-like fragility curve:

$$\widehat{\mathsf{Frag}}(IM;\, \delta_o) = \mathbb{P}\left[\log \Delta \geq \log \delta_o\right] = 1 - \mathbb{P}\left[\log \Delta \leq \log \delta_o\right]$$
$$= \Phi\left(\frac{\log IM - \left(\log \delta_o - B\right)/A}{\zeta / A}\right).$$

## Maximum likelihood estimation (ML)

Shinozuka *et al.* (2000)

- Lognormal shape:

$$\widehat{\mathsf{Frag}}(IM;\, \delta_o) = \Phi\left(\frac{\log IM - \log \alpha}{\beta}\right)$$

- Estimation of $\alpha$ and $\beta$ by maximum likelihood for each $\delta_o$:

$$\mathcal{L}\left(\alpha,\, \beta,\, \{IM_i\}_{i=1}^{N}\right) = \prod_{IM_i:\, \Delta_i \geq \delta_o} \left[\widehat{\mathsf{Frag}}(IM_i;\, \delta_o)\right] \prod_{IM_i:\, \Delta_i < \delta_o} \left[1 - \widehat{\mathsf{Frag}}(IM_i;\, \delta_o)\right]$$

# Non parametric methods

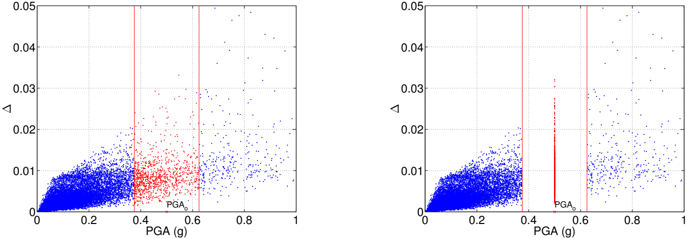
## Binned Monte Carlo estimate

Mai, Konakli & Sudret, Frontiers Struct. Civ. Eng., (2017)

- Suppose $N_s$ analyses are available for $IM = IM_o$, with $N_f$ such that $\Delta \geq \delta_o$. The fragility curve in this point could be estimated by Monte Carlo simulation:

$$\widehat{\text{Frag}}(IM_o) = \frac{N_f\,(IM_o)}{N_s\,(IM_o)}$$

- From the data cloud, a bin centered on $IM_o$ is considered, and points within the beam are "projected" onto the vertical line $IM = IM_o$ by linearization $\widetilde{\Delta_j}(IM_o) = \Delta_j \dfrac{IM_o}{IM_j}$.



# Kernel density estimation

## Fragility curves as a conditional CCDF

Mai *et al.* , Frontiers Struct. Civ. Eng., (2017)

$$\text{Frag}(a;\,\delta_o) = \mathbb{P}\left(\Delta \geq \delta_o | IM = a\right) = \int_{\delta_o}^{+\infty} f_\Delta(\delta | IM = a)\,\mathrm{d}\delta$$

where:

$$f_\Delta(\delta | IM = a) = \frac{f_{\Delta, IM}(\delta, a)}{f_{IM}(a)}$$

## Kernel density estimation

- The joint- and the marginal PDFs are estimated by:

$$\hat{f}_X\left(x\right) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \qquad \hat{f}_{\boldsymbol{X}}\left(\boldsymbol{x}\right) = \frac{1}{N\left|\boldsymbol{H}\right|^{1/2}}\sum_{i=1}^{N} K\left(\boldsymbol{H}^{-1/2}(\boldsymbol{x} - \boldsymbol{x}_i)\right)$$

NB: Use of a constant bandwidth in the logarithmic scale

Mai, C., Polynomial chaos expansions for uncertain dynamical systems – Applications in earthquake engineering, PhD Thesis, ETH Zurich, 2016

# Kernel density estimation

## Fragility curves as a conditional CCDF

Mai *et al.* , Frontiers Struct. Civ. Eng., (2017)

$$\text{Frag}(a;\,\delta_o) = \mathbb{P}\left(\Delta \geq \delta_o | IM = a\right) = \int_{\delta_o}^{+\infty} f_\Delta(\delta | IM = a)\,\mathrm{d}\delta$$

where:

$$f_\Delta(\delta | IM = a) = \frac{f_{\Delta, IM}(\delta, a)}{f_{IM}(a)}$$

## Kernel density estimation

- The joint- and the marginal PDFs are estimated by:

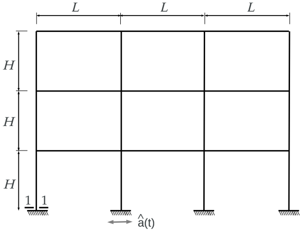
$$\hat{f}_X\left(x\right) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \qquad \hat{f}_{\boldsymbol{X}}\left(\boldsymbol{x}\right) = \frac{1}{N\left|\boldsymbol{H}\right|^{1/2}} \sum_{i=1}^{N} K\left(\boldsymbol{H}^{-1/2}(\boldsymbol{x} - \boldsymbol{x}_i)\right)$$

NB: Use of a constant bandwidth in the logarithmic scale

Mai, C., Polynomial chaos expansions for uncertain dynamical systems – Applications in earthquake engineering, PhD Thesis, ETH Zurich, 2016

# Outline

# Application : steel frame




- 2D steel frame submitted to synthetic ground motions
- Synthetic earthquakes generated in time domain

| Parameter | Distribution | Mean | Standard deviation | C.o.V |
|---|---|---|---|---|
| $f_y$ (MPa) | Lognormal | $264.2878$ | 18.5 | $0.07$ |
| $E_0$ (MPa) | Lognormal | $210000$ | 630 | $0.03$ |

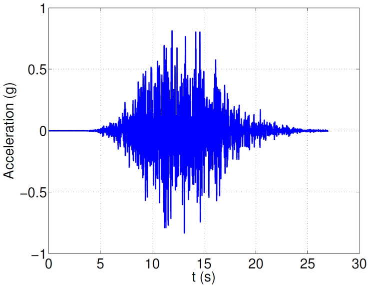
# Stochastic ground motion

## Stochastic excitation

- Obtained by a modulated filtered white noise process — Rezaeian & Der Kiureghian (2010)

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

- Parameters of the filter $\boldsymbol{\lambda} = \left(\omega_{mid}, \omega', \zeta_f\right)^{\mathsf{T}}$ are calibrated on recorded signals
- Global parameters (Arias intensity $I_a$, duration $D_{5-95}$, strong phase peak $t_{mid}$) are transformed into the parameters $\boldsymbol{\alpha}$ of the modulation function $q(t, \boldsymbol{\alpha})$ (e.g. gamma distribution)



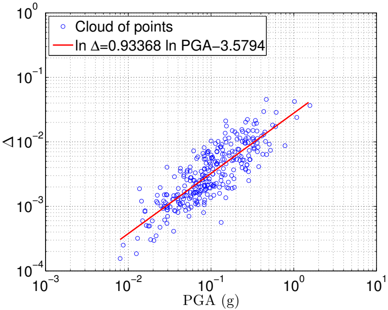
# Stochastic ground motion

## Parameters of the excitation

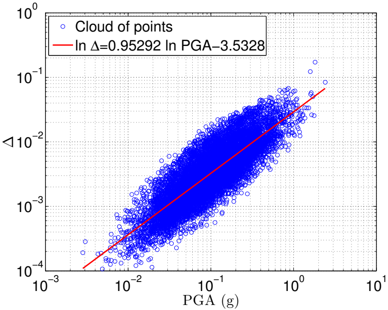
| Parameter | Distribution | Support | Mean | Standard deviation |
|---|---|---|---|---|
| $I_a$ (s.g) | Lognormal | $(0, +\infty)$ | 0.0468 | 0.164 |
| $D_{5-95}$ (s) | Beta | [5, 45] | 17.3 | 9.31 |
| $t_{mid}$ (s) | Beta | [0.5, 40] | 12.4 | 7.44 |
| $\omega_{mid}/2\pi$ (Hz) | Gamma | $(0, +\infty)$ | 5.87 | 3.11 |
| $\omega'/2\pi$ (Hz) | Two-sided exponential | [-2, 0.5] | -0.089 | 0.185 |
| $\zeta_f$ (.) | Beta | [0.02, 1] | 0.213 | 0.143 |

# Training and validation data

- Reference solution: Monte Carlo sampling of 10,000 non linear transient analyses
- PC-NARX: 300 samples



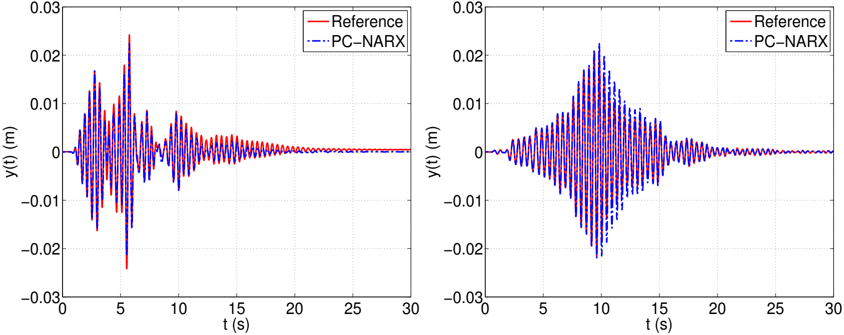

**300 simulations (training PC-NARX)**



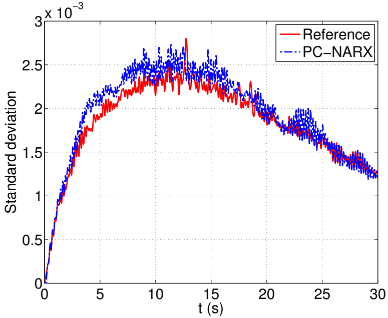

**10,000 simulations**

# Two trajectories (first floor displacement)

- Reference solution: Monte Carlo sampling of 10,000 non linear transient analyses
- PC-NARX: 300 samples



# Statistics of the first floor drift

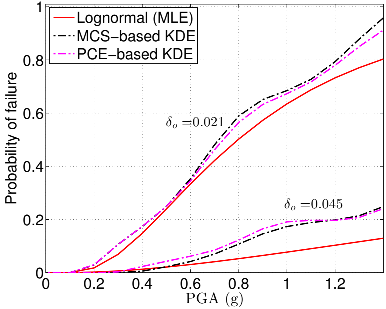



**Standard deviation** **PDF of max. drift**

# Fragility curve – maximal drift

- Reference solution: Monte Carlo sampling of 10,000 non linear transient analyses
- PC-NARX: 300 samples

## Fragility curves

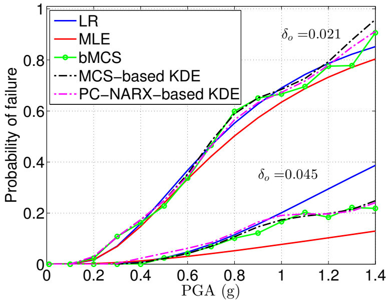



**Kernel density estimation**

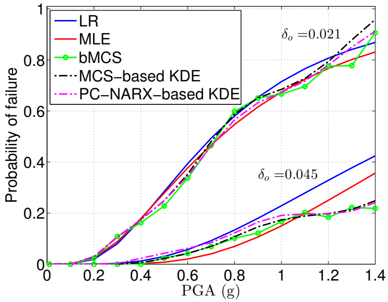
# Fragility curve – maximal drift

- Reference solution: Monte Carlo sampling of 10,000 non linear transient analyses
- PC-NARX: 300 samples

## Fragility curves




**Using 300 points**




**Using 10,000 points**

# Conclusions

- Polynomial chaos expansions are facing challenging issues when modelling time-dependent systems such as arising in structural dynamics
- A non-intrusive approach based on NARX models (from structural identification) and sparse PCE is proposed
- The accuracy is remarkable on the statistical moments (mean/std. deviation), PDF of the maximum output, but also on particular trajectories
- The method was successfully used for computing fragility curves in earthquake engineering applications

# Questions ?

**Chair of Risk, Safety & Uncertainty Quantification**

www.rsuq.ethz.ch

UQLab

www.uqlab.com

Thank you very much for your attention !

# References


M Shinozuka, M Feng, J Lee, and T Naganuma.
Statistical analysis of fragility curves.
*J. Eng. Mech.*, 126(12):1224–1231, 2000.

B. Ellingwood.
Earthquake risk assessment of building structures.
*Reliab. Eng. Sys. Safety*, 74(3):251–262, 2001.

S. Rezaeian and A. Der Kiureghian.
Simulation of synthetic ground motions for specified earthquake and site characteristics.
*Earthq. Eng. Struct. Dyn.*, 39(10):1155–1180, 2010.

C. V. Mai.
*Polynomial chaos expansions for uncertain dynamical systems – Applications in earthquake engineering*.
PhD thesis, ETH Zürich, Switzerland, 2016.

C.-V. Mai, M. D. Spiridonakos, E.N. Chatzi, and B. Sudret.
Surrogate modeling for stochastic dynamical systems by combining nonlinear autoregressive with exogeneous input models and polynomial chaos expansions.
*Int. J. Uncer. Quant.*, 6(4):313–339, 2016.

C.-V. Mai, K. Konakli, and B. Sudret.
Seismic fragility curves for structures using non-parametric representations.
*Frontiers Struct. Civ. Eng.*, 11(2):169–186, 2017.